分类号：　　学校代码：10389

密　级：　　学　　号：5221139005

福建农林大学

硕士学位论文

（专业学位）

基于机器学习与区块链的 SDN 控制器 DDoS 防御策略研究

学习形式：全日制

学位类别：电子信息硕士专业学位

专业领域：电子信息

研究方向：网络与信息安全

学生姓名：田佳霖

指导教师：舒兆港 副教授

完成时间：二〇二五年五月

DDoS Defense Strategy Based on Blockchain and Unsupervised Learning Techniques in SDN

By

Jialin Tian

Supervised by Associate Professor Zhaogang Shu

A Thesis Submitted to

Fujian Agriculture and Forestry University

in Partial Fulfillment of the Requirements

for

Professional Master's Degree in Electronic Information

College of Computer and Information Science

Fujian Agriculture and Forestry University

Fujian, P.R. China

Completion Date (May 2025)

Commencement Date (May 2025)

独创性声明

本人声明，所呈交的学位（毕业）论文，是本人在指导教师的指导下独立完成的研究成果，并且是自己撰写的。尽我所知，除了文中作了标注和致谢中已作了答谢的地方外，论文中不包含其他人发表或撰写过的研究成果。与我一同对本研究做出贡献的同志，都在论文中作了明确的说明并表示了谢意，如被查有侵犯他人知识产权的行为，由本人承担应有的责任。

学位（毕业）论文作者亲笔签名：王佳秉　日期：2025年5月23日

论文使用授权的说明

本人完全了解福建农林大学有关保留、使用学位（毕业）论文的规定，即学校有权送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

保密，在　　年后解密可适用本授权书。 □

不保密，本论文属于不保密。 ☑

学位（毕业）论文作者亲笔签名：王佳秉　日期：2025年5月23日

指导教师亲笔签名：　　日期：2025年5月23日

目 录

图 录

表 录

摘 要

软件定义网络（Software Defined Networking, SDN）自身具有集中式控制的架构，从而导致其 SDN 控制器逐渐成为了分布式拒绝服务（Distributed Denial of Service，DDoS）主要攻击的目标之一。一旦攻击者利用恶意流量占用控制器的计算与存储资源，网络将面临严重的服务中断。现有的 DDoS 防御方法在 SDN 环境下难以有效应对攻击的动态性与隐蔽性，面临高误报率、实时检测能力不足等问题。为此，本文提出一种基于区块链与机器学习的 DDoS 防御策略，利用区块链的去中心化和不可篡改特性增强流表管理的安全性，同时结合无监督学习算法，实现对异常流量的精准检测与智能缓解，以提高 SDN 网络的安全性和鲁棒性。

在 SDN 架构中，流表是控制器管理和调度网络流量的关键数据结构，但 DDoS 攻击往往通过大量伪造流请求迅速耗尽流表资源，导致合法流量被拒绝，从而加剧控制器的负担。本文结合现有对 SDN 流表保护和可疑攻击源的检测两方面的研究进展，提出了一种新的能够在 SDN 网络环境下检测到 DDoS 攻击并缓解其攻击的双层策略。本文的主要研究内容与创新点如下：

（1）在 SDN 流表保护方面，我们将 SDN 的流表存储容器与控制器进行剥离，使之成为独立的流表数据存储层。在 SDN 控制器的北向接口处，使用了联盟链 FISCO BCOS 区块链作为 SDN 存储流表的容器，利用区块链的去中心化和不可篡改特性，防止攻击者篡改、伪造流表数据。然后，设计了一种基于智能合约的跨域数据流传输方案，以确保流表在高频访问下的正常转发，同时减少因 DDoS 攻击导致的异常流表更新和存储污染。实验结果表明本方案能有效提升流表保护安全性，且在小规模的 SDN 网络拓扑中本方案能显著增强 SDN 控制器的流表转发的稳定性，这为后续实时读取流表中数据进行可疑攻击源检测奠定基础。

（2）在可疑攻击源的检测方面，将 DDoS 攻击检测看成是一个二元分类问题，把 SDN 网络中的数据传输分为良性的网络流与恶意网络流两种类别，提出了一种使用了随机森林（Random Fores，RFA）算法，互信息（Mutual Information，MI）算法，以及递归特征消除（Recursive Feature Elimination，RFE）算法组成的复合特征值筛选方法对数据进行相关性系数计算与相关性排序，与数据筛选。为尽可能模拟实际情景下，作为攻击源的数据，我们使用了 CIC-DDoS2019 公共的 DDoS 数据集作为数据支撑，其中包含 11 种不同类型 DDoS 攻击的 77 个基于网络流的统计特征。经过筛选与优化，我们最终选取了其中 7 个关键特征用于实时分析，显著提升了检测的精准度与效率。

（3）对于恶意的网络流，进一步将其分成了易检测和不易检测两部分，并提出了一种双层防御方法，第一层结合时间序列和令牌桶两种算法通过对网络流访问的频率进行监控，防御易检测的攻击，第二层使用了无监督机器学习算法中的孤立森林算法，通过先前对区块链中存储的数据流进行数据筛选后确定的 7 个高相关性数据特征进行异常值计算，来判定

不易检测的攻击。相比于同类型防御策略，DDoS 攻击识别的准确率达到 97.66%。与其他同类方法比较，在判定准确度相近的情况下，在小规模的 SDN 网络拓扑中部署本策略对 DDoS 攻击具有更快的判定速度。

关键词：软件定义网络；分布式服务攻击；区块链；机器学习

ABSTRACT

Software Defined Networking (SDN), due to its centralized control architecture, has increasingly made the SDN controller a primary target for Distributed Denial of Service (DDoS) attacks. Once attackers exploit malicious traffic to exhaust the controller's computational and storage resources, the network may suffer severe service disruptions. Existing DDoS defense methods in SDN environments struggle to effectively cope with the dynamic and stealthy nature of such attacks, often facing high false positive rates and insufficient real-time detection capabilities. To address these challenges, this paper proposes a DDoS defense strategy based on blockchain and machine learning. By leveraging the decentralization and immutability features of blockchain to enhance the security of flow table management, and integrating unsupervised learning algorithms to accurately detect and intelligently mitigate anomalous traffic, this strategy aims to improve the security and robustness of SDN networks.

In SDN architectures, flow tables serve as critical data structures for managing and directing network traffic. However, DDoS attacks often exploit large-scale forged flow requests to rapidly exhaust flow table resources, causing legitimate traffic to be rejected and further overburdening the controller. To counter this, this paper integrates research advancements in SDN flow table protection and suspicious attack source detection, proposing a novel dual-layer strategy capable of detecting and mitigating DDoS attacks in SDN environments. The main contributions of this paper are as follows:

The flow table storage container is decoupled from the SDN controller and implemented as an independent storage layer. At the northbound interface of the SDN controller, the consortium blockchain FISCO BCOS is employed as a secure container for flow table storage. The decentralized and tamper-resistant properties of blockchain ensure that attackers cannot modify or forge flow table data. Furthermore, a cross-domain data transmission scheme based on smart contracts is designed to maintain normal flow table forwarding under high-frequency access while reducing abnormal updates and storage contamination caused by DDoS attacks. Experimental results demonstrate that this approach significantly enhances the security of flow table protection and improves flow table forwarding stability in small- and medium-scale SDN network topologies. This lays a solid foundation for subsequent real-time flow table analysis for attack source detection.

The detection of DDoS attacks is formulated as a binary classification problem, where network traffic in the SDN environment is categorized as either benign or malicious. A hybrid feature selection method, combining Random Forest (RF), Mutual Information (MI), and Recursive Feature Elimination (RFE), is proposed to compute correlation coefficients, rank features, and filter data accordingly. To simulate real-world attack scenarios, this study utilizes the publicly available CIC-DDoS2019 dataset, which contains 77 network flow-based statistical features representing 11 different types of DDoS attacks. After feature selection and optimization, seven key features are identified for real-time analysis, enhancing detection accuracy and efficiency.

For malicious network traffic, it is further divided into easily detectable and hard-to-detect categories. A two-layer defense strategy is proposed. The first layer integrates time series analysis and the token bucket algorithm to monitor network access frequency and defend against easily detectable attacks. The second layer employs the Isolation Forest algorithm, an unsupervised machine learning method, to calculate outliers based on six highly correlated features extracted from blockchain-stored network flow data, enabling the identification of hard-to-detect attacks.

Compared to similar defense strategies, the proposed approach achieves a DDoS detection accuracy of 97.66%. Furthermore, under comparable detection accuracy, this strategy demonstrates faster attack identification in small-scale SDN network topologies.

KEY WORDS: Software Defined Networks; Distributed Denial of Service; Blockchain; Machine Learning

1 绪论

1.1 研究背景及意义

在互联网技术日新月异的今天，网络架构的灵活性和效率需求日益凸显。根据 2024 年 3 月发布的第 53 次《中国互联网络发展状况统计报告》数据，截至前年 12 月，中国网民规模达 10.92 亿，互联网普及率为 77.5%[1]。我国互联网在过去三十年间取得了跨越式的发展，已建成全球规模最大、技术领先的互联网基础设施。报告显示互联网在中国社会中已经深度渗透。随着数字消费、生成式人工智能等新兴应用的快速发展，网络流量、业务需求不断增加，这些均给网络运营商和服务提供商带来了巨大的挑战。

为了有效应对当前网络领域的重大变革，软件定义网络[2]由此得以诞生。这种网络架构是基于数据包转发平面与控制平面的分割而设计的。该架构拓展了网络设计和管理的可能性，它将网络控制所需的一系列操作都集中在了一个被叫做 SDN 控制器的软件组件中。通过使用集中控制的方式，大大简化了执行新服务、选择管理策略以及从软件层面重新配置网络的要求。在 SDN 网络拓扑中，控制平面与 SDN 控制器之间通过北向接口相互连接。控制平面中的上层应用可对 SDN 控制器发送命令和请求。SDN 控制器与数据平面之间则通过南向接口来实现控制器对平面内网络设备发送调控的指令，同时接收来自网络设备的各种状态信息。这种由 SDN 构建的网络架构，是一种可编程网络[3]。目前，已经有许多种基于 SDN 研发的成熟网络协议。Openflow 协议就是其中常见的协议中的一种[4]。

软件定义网络作为一种新兴的网络架构，它能够在一定程度上缓解传统网络所面临的流量与管理压力，但与此同时，也不可避免地面临多种网络攻击的威胁。DDoS 攻击是通过发送大量恶意数据包，并通过分布在各处的受损节点耗尽可用资源[5]来阻止合法用户的访问。通过研究发现，SDN 的数控分离模型可以让用户和上层应用对网络有更多控制权[6]，但仍会面临到 DDoS 攻击的威胁。DDoS 攻击通过耗尽 SDN 拓扑中交换机流表缓存或控制器的资源，或阻塞其拓扑中的链路带宽来攻击 SDN 网络。DDoS 攻击对网络运营商和互联网服务提供商的威胁与日俱增，其也被认为是网络安全领域最棘手的问题之一，目前还没有完美的解决方案[7]。特别是经过 5G 时代的发展，网络中设备的数量呈指数级增长，对防御 DDoS 攻击构成巨大威胁。因此，研究如何在 SDN 中防御 DDoS 也具有重要意义。

为了用户能够体验到更为优质的网络服务，面对当下愈加复杂的网略环境，网络流量监管技术被引入至各类系统之中，旨在实现对潜在恶意流量的及时感知

与准确识别。所谓恶意流量检测[8]，通常是指在网络中对流量数据进行采集，结合特定的技术手段构建分类模型，对数据流进行判别和归类，从而识别出潜藏的异常行为。然而，受限于实际网络环境的复杂性以及恶意流量本身低频、分布稀疏的特征，研究中往往难以获得数量充足、代表性强的训练样本[9]，这在一定程度上制约了检测方法的有效性与泛化能力。近年来，如何进行小样本条件下的恶意流量检测问题引发了广泛讨论，研究人员普遍使用机器学习技术[10]和深度学习技术[11]对其进行研究，以期在有限数据条件下提升恶意流量的检测能力。相比与深度学习技术需要事先对数据添加标签的繁琐步骤，且易于过拟合，机器学习技术能够更快的识别出恶意流量，内存消耗更小，因此在 SDN 中异常流量检测方面具有较好的应用前景。

已有 SDN 防御 DDoS 技术具有一定的局限性，表现为：一方面传统 DDoS 防御策略通常依赖流量过滤与流量识别，易受攻击流量的波动性较大，造成误报率高，防御效果不够稳定；另一方面，这类方法一般很难处理 DDoS 攻击高度动态性与分布式的特点，很难对隐蔽攻击进行实时准确的识别与抵御。SDN 环境中攻击者可通过迅速消耗控制器资源或者破坏流表数据来造成全网瘫痪。传统 DDoS 防御方法在应对上述挑战时逐渐显露出自身防护能力的欠缺，迫切需要引进更有效和更灵活的防御手段。

在此背景下，区块链技术的广泛应用为解决 SDN 网络拓扑的域间协作提供了新的思路[12]。区块链技术[13]作为结合了一系列现有的技术，例如密码学、Merkle 树和共识机制，它具有去中心化能力、数据安全保护、数据共享功能和分布式特性，这些优势为 DDoS 攻击的协同防御提供了坚实的基础[14]。区块链技术解决了多个域之间的数据共享以及相互合作的域之间的信任问题[15]。智能合约[16]被用于描述“由各方执行的协议”。它将参与制定合约的各方的协议编写成规则代码。它是一种运行在区块链上的可重用、模块化、自动执行的脚本。区块链技术凭借其智能合约机制与内建的共识协议，在解决跨域场景下的信息交互与协同难题方面展现出显著优势，能够在缺乏信任基础的网络环境中实现去中心化的协作机制。此外，以比特币、以太坊及分布式账本为代表的区块链系统已在多个实际应用领域中验证了其在数据安全性与透明性方面的优越性能，有效提升了系统整体的可信度和防篡改能力，在 SDN 网络中应用于 DDoS 防御也有很大发展潜力[17]。

1.2 国内外研究现状

针对 SDN 网络拓扑的 DDoS 防御领域，现有工作中主要集中于，流表数据存储和可疑攻击源的检测两方面。这两点之中，对于可疑攻击源检测多使用学习算法来进行，包含机器学习算法和深度学习算法。通过对数据集进行分析，得出

可疑攻击地址进行防御，这类方法往往因使用算法不同而具有多样性。而对 SDN 流表保护的要求，有别于传统的信息存储思路，伴随近几年区块链技术的火热发展，大家开始尝试利用区块链的高安全性的属性来弥补传统网络信息存储方案的不足[18]。其中又经常用其来存储攻击源地址的黑白名单，这种方式虽然可以有效阻止对攻击源数据的转发，但仍可能存在着主观设定问题。Amezcua Valdovinos 等人[19]介绍了 SDN 网络架构并且概述了应对 DDoS 攻击，在 SDN 领域的研究工作和需要解决的问题。从该综述文章内容可知，继承于对传统 DDoS 攻击的防御思路，SDN 网络的 DDoS 防御策略也拥有多个角度，这里将其归纳为数据存储管理和攻击源检测与防御两部分。

综上所述，可以得出在 SDN 防御 DDoS 攻击领域，我们可以采取从 SDN 流表保护和 DDoS 攻击检测两个方向进行防御措施的设计。目前该领域研究已有一定成果，但在平衡检测准确性、实时性以及可扩展性上仍然有待进一步深入。伴随着区块链、机器学习等技术的进一步发展，将区块链技术与机器学习算法结合的策略设计，会更有效的考虑到上述三点问题，从而提升 SDN 网络在面对复杂 DDoS 攻击时的整体防御能力。

1.2.1 SDN 流表保护

从提升 SDN 流表保护的角度，结合时下发展迅速的区块链技术，可以有效扩展传统的存储网络信息保护思路。近年来，随着区块链技术的广泛应用，多个研究已尝试将其引入 SDN 和网络安全领域，以解决数据共享、存储安全及隐私保护等问题。Gao 等人[20]提出了一种基于区块链的数据共享框架，结合 SDN 和代理重加密（Proxy Re-Encryption, PRE）技术，旨在提升工业物联网环境下边缘计算的安全性和隐私保护。该框架利用区块链的去中心化和不可篡改特性，确保工业物联网设备授权管理和数据共享过程的安全性，结合 PRE 技术，实现数据的细粒度访问控制。虽然该方案在提升数据安全性和隐私保护方面有明显优势，但未深入探讨 SDN 网络中的流表保护与 DDoS 防御的具体问题，这与本文提出的基于区块链的 DDoS 防御策略相比，应用场景不同，研究重点有所差异。Núñez-Gómez 等人[21]提出了一种增强的分布式资源分配和访问控制框架 S-HIDRA，结合区块链的智能合约功能，实现了去中心化、自主且可靠的资源编排。该框架在 SDN 中集成区块链，能够动态管理网络服务的连接，并提高网络服务的高可用性和低延迟。该框架提高了 SDN 网络资源管理的效率，但其侧重于服务管理和资源编排，未专门针对 DDoS 攻击防御进行优化。本文则更专注于在 SDN 网络中有效提高了对 DDoS 攻击的防护能力。

Zeng 等人[22]提出了一种利用区块链建立全局信任关系的方案，设计了信誉度评价机制，以确保多域 SDN 支持的物联网网络中的安全可路由性。该方法通

过在区块链上记录和管理各节点的信誉度信息，构建了一个透明且不可篡改的信任体系。然而研究也指出，信誉值的定义可能存在主观性，影响评价机制的客观性。这种基于信誉的信任体系对提高网络安全性作用较为有限。本研究则针对DDoS攻击的高动态性和隐蔽性，提出了基于区块链和机器学习的双层防御策略，更具针对性地解决了 DDoS 攻击带来的流表资源耗尽问题。Hang 等人[23]提出了一种基于区块链的 SDN 交换机信息安全保护系统，旨在解决现代农业 4.0 环境中的安全问题。该系统通过区块链确保数据的机密性、隐私性和可扩展性，增强了信息传输的安全性和可靠性。实验表明，该系统在面对 DDoS 攻击时，通过 SDN 控制器实现的 DDoS 缓解程序，能够有效提高系统稳定性和数据完整性。该研究的优势在于能够通过 SDN 控制器自动缓解 DDoS 攻击，其应用场景主要针对智能农业领域，具体解决的是农业环境中的网络安全问题，本文通过将区块链技术与流表保护和 DDoS 攻击缓解结合，提出了一种更为普遍适用的解决策略。任等人[24]提出了一种基于区块链的多控制器 SDN 安全认证机制，旨在提升软件定义物联网环境下的网络安全性。该框架结合区块链技术，通过去中心化和不可篡改的特性，解决了传统 SDN 架构中单点故障和多控制器之间的数据一致性问题。通过引入工作量证明算法和声誉机制，确保控制器之间的数据视图一致，避免了伪造和欺骗攻击的风险。虽然该研究对 SDN 网络的多控制器一致性问题提供了创新的解决方案，但它更侧重于控制器之间的数据一致性和认证机制，对于如何针对 DDoS 攻击进行高效防御的研究较少。

综上所述，面对当下针对 SDN 网络的 DDoS 攻击，我们除了采用传统网络的防御措施外，也需要针对 SDN 网络的特点设计专门的防御措施。当前研究领域大多专注于 SDN 流表保护或可疑攻击源的有效识别两方面研究，但在实际部署中，又各自存在着诸如检测的准确性、响应的实时性以及策略的可扩展性较差等问题。而本文提出的将区块链的去中心化机制与机器学习的自适应特性相结合，不仅是一种SDN防御DDoS攻击的新思路，而且也能更好的平衡上述三点问题，实现有效提升 SDN 网络整体防御 DDoS 攻击能力的期望。

1.2.2 可疑攻击源的检测

对于针对可疑的DDoS攻击源检测方面，已有研究提出了多种有效的检测方法。Ma 等人[25]，提出了一种结合异构集成特征选择和随机森林算法的 DDoS 攻击实时检测方法。该方法首先通过异构集成特征选择，从大量特征中筛选出最具判别力的特征子集，然后利用随机森林算法构建分类模型，对网络流量进行实时监测和分类。实验结果显示，该方法在检测精度和速度方面均有良好表现。然而，研究也指出，随机森林作为一种监督学习算法，对计算资源的需求较高，这在资源受限的环境中可能成为限制因素。与之相比，本文结合区块链技术来增强流表

的安全性，并使用无监督学习算法进行攻击源检测，不依赖于大量标注数据，且能够减少计算资源的消耗，特别适用于 SDN 网络中对实时性要求较高的防御任务。Marvi 等人[26]，提出了一种结合特征选择和特征提取的混合方法，利用无监督机器学习技术检测不同类型的 DDoS 攻击。该方法通过特征选择减少数据维度，提取关键特征，然后采用无监督学习算法对网络流量进行聚类分析，识别异常模式。这种方法无需预先标记的数据，适用于未知攻击类型的检测，但实验也指出该方法在复杂网络环境中的适用性可能受到限制瓶颈。相较而言，本文提出的防御策略，结合了无监督学习的灵活性，能够有效应对大规模网络中的异常流量检测和攻击源识别问题。谢等人[27]提出了一种基于强化学习的 SDN 安全防护机制，旨在提升 SDN 环境下的网络安全性。通过引入强化学习算法，该机制能够动态地识别和应对网络中的各种安全威胁，实现对网络流量的智能监控和异常检测。该方法在提高 SDN 网络的安全性和自适应能力方面具有显著优势。但强化学习模型的训练过程可能需要大量的计算资源和时间，这在实际应用中可能导致效率低下。

相比之下，本文的方案避免了强化学习的训练，能够在有限的计算资源条件下实现高效的 DDoS 防御。杨等人[28]，在其研究中提出了一种基于 Rényi 熵与随机森林结合极限梯度提升（eXtreme Gradient Boosting，XGBoost）算法的 DDoS 攻击检测方法，SDN 环境设计。该方案通过使用 Rényi 熵对流量特征进行提取，并结合 RF 和 XGBoost，运用集成学习增强了攻击检测模型的准确性和效率。该模型的核心优势在于能够动态适应 SDN 环境下的流量变化。但由于其较高的计算需求，尤其在大规模 SDN 网络中，可能会面临计算资源不足的问题。与此相比，本文通过减少对计算资源的依赖，尤其是在流量分析和特征选择的过程中，能够更好地应对 SDN 网络中流表安全和 DDoS 防御的挑战。王等人[29]，在其研究中提出了一种基于深度学习的 SDN 异常流量分布式检测方法。该方法通过在云端服务器部署判别器，并在各个 SDN 控制器上部署生成器，构建了一个分布式生成对抗网络（Distributed Variational Autoencoder Wasserstein Generative Adversarial Network，D-VAE-WGAN）。其主要目标是解决传统集中式异常流量检测方法在处理大规模 SDN 时的计算开销过大、链路拥塞以及单点故障的问题。然而由于该方法基于深度学习，计算资源的需求较高，尤其是在训练阶段。与此不同，本文采用机器学习算法，充分考虑了计算资源的优化，避免了深度学习方法的高计算开销。

综合来看，现有的可疑 DDoS 攻击源检测方法虽然在不同程度上提升了检测精度和效率，但大多数方法仍面临计算资源需求高、实时性不足等问题，尤其在大规模 SDN 网络环境中应用时，如何平衡检测精度、计算开销仍然是一个重要挑战。本文通过结合机器学习算法，提出一种高效、节省资源的 DDoS 防御策略，

针对性地提高了 DDoS 攻击源的检测能力，平衡了衡检测精度与计算开销，具有较高的应用前景。

1.2.3 流表保护与攻击源检测相结合

除了在 SDN 网络中单独关注流表保护和攻击源检测之外，近年来也有一些研究尝试将这两部分相结合，探索如何通过结合区块链与机器学习技术进行 DDoS 防御。Lian 等人[30]提出了 FRChain，这是一种基于区块链的 SDN 数据转发安全方案。FRChain 通过在区块链中存储流规则来确保 SDN 数据转发的安全性，同时利用区块链的去中心化特性提供了流量控制的透明性和不可篡改性。该方案采用了节点投票机制来检测网络中的可疑节点，通过多个节点的投票来判断节点的可信度。尽管该方案在提高数据转发安全性方面具有优势，但投票过程需要较长时间，这使得该方法不能即时响应攻击，从而影响了其在实时攻击检测中的应用。相比之下，本文提出的基于区块链和机器学习的双层防御策略，能够在实时检测和流表保护方面提供更高效的防御响应，特别是在 SDN 环境下，对 DDoS 攻击的快速响应能力尤为关键。J Jian 等人[31]提出了一种基于信息论的 DDoS 检测方法，并结合区块链中的智能合约用于网络安全防御。该方法通过信息论中的熵值计算来判断网络流量的正常与异常，并使用区块链技术记录黑白名单，以实现自动化的攻击防御。智能合约能够自动化执行攻击识别和响应任务，但由于区块链的黑白名单机制可能带有一定的主观设定，因此在实际应用中可能存在一定的偏差和局限性。相比之下，本文提出的方案通过结合无监督学习算法进行攻击源检测，并借助区块链存储流表信息，在保障流表数据安全性的同时，确保了防御机制的高效性和实时性，避免了传统方法中可能存在的性能瓶颈。Li 等人[32]提出了 BlockCSDN 框架，旨在通过区块链技术实现 SDN 网络中的信息管理和入侵检测。BlockCSDN 框架结合区块链的透明性和去中心化特性，在网络中实施协作式入侵检测，利用区块链记录黑白名单来实现对网络节点的管理。然而，区块链的黑白名单机制在一定程度上仍受到主观设定的影响，且该方案的实时性可能受到区块链处理速度的制约。宋等人[33]，提出了一种基于机器学习的 DDoS 攻击检测方法，专为区块链环境设计。该方法采用了 LightGBM 算法，以解决在区块链节点中部署传统 DDoS 攻击检测系统时遇到的计算资源消耗过大、部署困难的问题。通过引入多元网络流量特征提取模型，并结合特征相关注意力机制与多层感知机（MLP），研究旨在提升 DDoS 攻击检测的准确性与效率。杨等人[34]，提出了一种新的跨域协同 DDoS 攻击检测方案，结合 SDN 与区块链技术增强信息协同与防御能力。该方法使用基于熵值的异常检测方法进行流量的宏观检测，并通过改进的支持向量机（SVM）进一步分析潜在攻击流量。虽然该方法具有较低的计算负荷和较高的适应性，但基于熵值的检测方法可能在低速攻击

流量识别上存在局限，且区块链的性能可能成为大规模流量下的瓶颈。相比之下，本文提出的方案结合了区块链的透明性和不可篡改性，能够有效解决实时性问题，同时通过机器学习算法提高攻击源检测的准确性，克服了基于熵值方法在低速攻击流量识别中的不足。

总体来看，当前的 SDN 网络 DDoS 防御研究主要集中于流表保护和攻击源检测的分开处理，尽管有一些尝试将这两部分结合，但大多数方法仍未能有效克服区块链在实时性和效率上的固有局限。例如，较低的读写效率和处理速度等问题仍然是区块链技术在 DDoS 防御中应用的瓶颈，尤其是在面对大规模 DDoS 攻击时，区块链的性能可能成为阻碍其实时检测和防御效果的关键因素。为了解决这些问题，本文提出了一种创新的解决方案，首先在 SDN 网络中引入了区块链存储流表信息，并通过设计高效的数据结构与智能合约优化了流表数据的存取速度。与传统流表存储方法相比，该方案减少了对 SDN 控制器计算资源的需求，提高了流表更新的实时性与准确性。其次，通过结合机器学习算法，本文提出了一种基于无监督学习的 DDoS 攻击源实时检测方法，能够动态识别异常流量并根据检测结果触发防御机制。该方法避免了传统基于特征的检测方法中的误报问题，提升了对隐蔽攻击源的识别能力，从而在减少计算负担的同时，增强了防御的精确性和效率。通过这些改进，本文的方案显著提高了 SDN 网络在面对大规模 DDoS 攻击时的防御能力和响应速度。

1.3 本文主要贡献

1.3.1 研究目标

由上文可见 SDN 网络的 DDoS 攻击防御研究领域中，现有研究工作大多集中于 SDN 流表保护和可疑攻击源的检测这两点上。但大多在各自方向存在问题，且专注于单一方向研究。本研究的主要目标是设计并实现一种基于区块链技术和无监督学习方法的 DDoS 攻击检测与防御策略，专门针对软件定义网络环境中的 DDoS 攻击防御问题。具体目标包括：

1、提出 SDN 控制器针对 DDoS 攻击的双层防御策略：结合区块链技术与无监督学习算法，在 SDN 网络中实现高效的 DDoS 攻击检测与防御。通过智能合约和区块链去中心化的特性，提升流表存储安全性并确保数据流的安全，保障网络的稳定性和抗攻击能力。

2、提升 SDN 流表安全性：将 SDN 流表信息存储在区块链上，利用区块链技术的不可篡改和去中心化特性增强流表数据的安全性。设计并应用智能合约处理流表的增、删、改操作，确保流表信息的安全存储与实时更新，从而防止 DDoS

攻击通过篡改或破坏流表规则来实施。

3、流表数据特征筛选：在 DDoS 攻击检测过程中，利用数据特征筛选方法进行流量数据的特征选择和优化。通过筛选最具判别性的流量特征，减少计算负担，提高 DDoS 攻击检测的准确性和实时性。

4、加强可疑攻击源检测：采用双层检测机制，通过结合基于规则的检测和机器学习算法，动态识别和检测 SDN 网络中的 DDoS 攻击。第一层通过基于流表的规则检测明显的攻击流量，第二层结合无监督学习算法（如隔离森林）进行深度检测，识别隐蔽或变化较大的 DDoS 攻击。

通过这些目标，本研究将为 SDN 环境中的 DDoS 防御提供创新的解决方案，结合区块链和机器学习算法的优势，在提升流表保护、攻击源检测和防御效率的同时，实现更高效、可靠的 DDoS 防御系统。

1.3.2 研究内容

本研究围绕 DDoS 攻击防御展开，主要包含以下几个研究内容：

1、基于区块链的流表安全存储：通过 FISCO BCOS 联盟链作为 SDN 控制器北向接口流表数据存储容器，通过智能合约对流表信息进行处理和管理。该方法不仅保障了数据存储的安全性，还通过高频访问要求的支持，使得区块链在 SDN 网络中得以有效应用，为防御 DDoS 攻击提供了新的思路。

2、双层防御策略设计：提出了一种基于区块链技术和无监督学习算法的双层 DDoS 防御策略。该策略通过将区块链技术与 SDN 网络的流表数据存储结合，提升了数据存储的安全性，并通过时间窗算法和令牌桶算法监控和限制易检测的 DDoS 攻击。同时，使用无监督学习中的 Isolation Forest 算法，检测那些较难察觉的 DDoS 攻击。

3、数据筛选优化：在 DDoS 攻击检测中引入了复合特征选择算法，该算法结合了递归特征消除、随机森林和互信息等方法，以提高特征选择的准确性。在此基础上，结合 Isolation Forest 算法进行二分类检测，从而有效区分正常流量与异常流量，并显著提高了 DDoS 攻击检测的准确性和稳定性。

1.4 论文结构和章节安排

根据 DDoS 防御策略问题的研究现状和发展趋势，本文着重研究基于机器学习和区块链技术的 DDoS 防御策略设计。因此，本文结合机器学习算法和区块链技术，提出了一种新的防御策略，并在此基础上针对各项指标进行优化。使之和同类 DDoS 防御策略相比，获得了不错的检测与防御性能。本文的研究内容和各

个章节安排如下图 1-1。

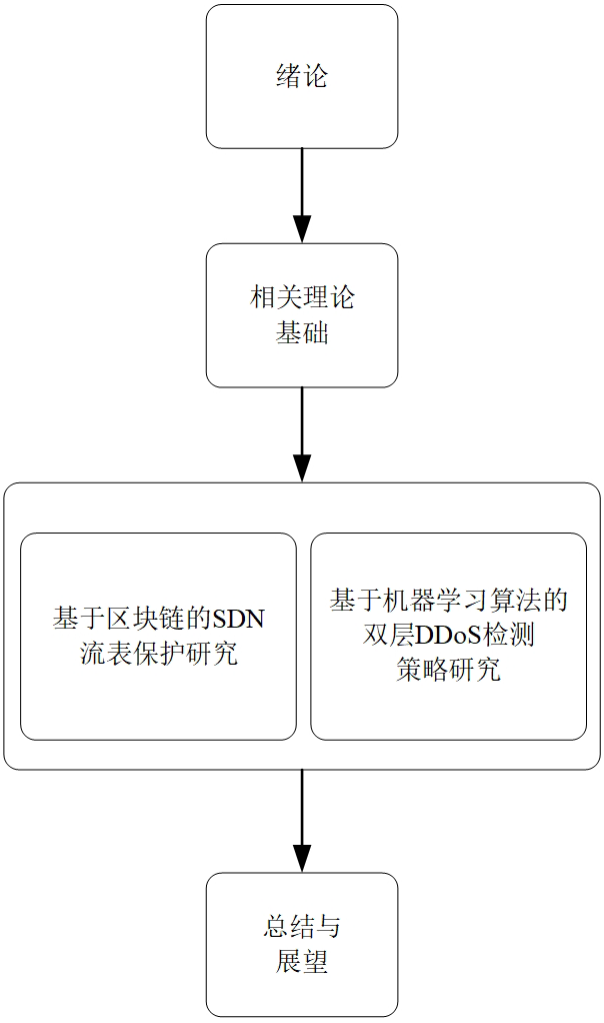

图 1-1 论文整体结构图

第一章绪论首先介绍 DDoS 攻击对 SDN 网络的安全威胁，分析当前研究现状，指出现有防御策略的不足，并明确本文的研究目标。最后，简单阐述了主要论文贡献与全文结构。

第二章相关理论基础介绍 SDN 网络的基本架构、DDoS 攻击的分类与防御挑战、区块链的核心技术（如共识机制、智能合约等），以及机器学习算法（孤立森林、随机森林、互信息等）在异常检测中的应用，为后续研究奠定理论基础。

第三章基于区块链的 SDN 流表保护研究提出了一种基于区块链的 SDN 流表保护方案，通过 FISCO BCOS 联盟链存储流表数据，提高数据安全性，并通过智能合约实现高效访问。本章包括流表保护方案设计，将 SDN 流表存储在区块链上，确保流表数据的安全性和不可篡改性；实验验证，搭建 SDN+区块链测试环境，评估本方案的存储稳定性和数据访问效率，并与传统 SDN 流表存储方式进行对比分析。

第四章基于机器学习算法的双层 DDoS 检测策略研究提出了一种结合区块

链流表保护的双层 DDoS 检测策略。第一层检测基于时间窗口算法和令牌桶机制，快速筛选和抑制大规模 DDoS 攻击。第二层检测使用复合特征选择算法（互信息+随机森林+递归特征消除）优化特征提取，并通过孤立森林算法检测隐蔽 DDoS 攻击，提高检测精度。实验评估采用 CIC-DDoS2019 数据集，评估所提策略的检测精度和实际应用效果，并与现有 DDoS 防御方案进行对比分析。

第五章总结与展望总结本文的主要研究成果，包括区块链流表存储方案与双层 DDoS 检测策略，并提出未来研究方向，如优化区块链存储性能、提升检测算法实时性、扩展防御策略适用范围等。本论文结合区块链与机器学习技术，为SDN环境下的 DDoS 防御提供了一种创新性解决方案，并在实验中验证了其有效性，为未来 SDN 安全研究提供了重要参考。

2 相关理论基础

2.1 SDN 相关技术

SDN 架构是一种新一代的网络架构设计，其核心思想在于通过将网络的控制平面与数据平面进行功能分离，从而实现网络控制与转发职能的解耦。它的基本架构如图 2-1 所示。在 SDN 中，控制平面主要负责网络的策略制定与整体调度，而数据平面则专注于数据流的具体转发与处理操作。SDN 不仅简化了网络管理流程，还为新型网络协议的研发与功能扩展提供了更加开放的实验平台。作为一种区别于传统网络模式的新型网络架构，SDN 具有三大核心特性。

1、转发与控制分离：SDN 通过将控制平面与数据平面解耦，将协议计算和流表生成的任务交由控制器处理，而底层硬件转发设备仅负责数据包的转发操作。

2、集中式控制：网络中的网关设备由控制器统一管理和调度，控制器通过下发流表来实现网络控制，避免了对各硬件设备的逐一配置，从而大幅简化了网络管理流程。

3、支持可编程性：控制器为外部开发提供开放接口，允许网络管理员通过编程定义特定网络功能，并将其部署在控制器上运行[35]，以实现个性化的配置。

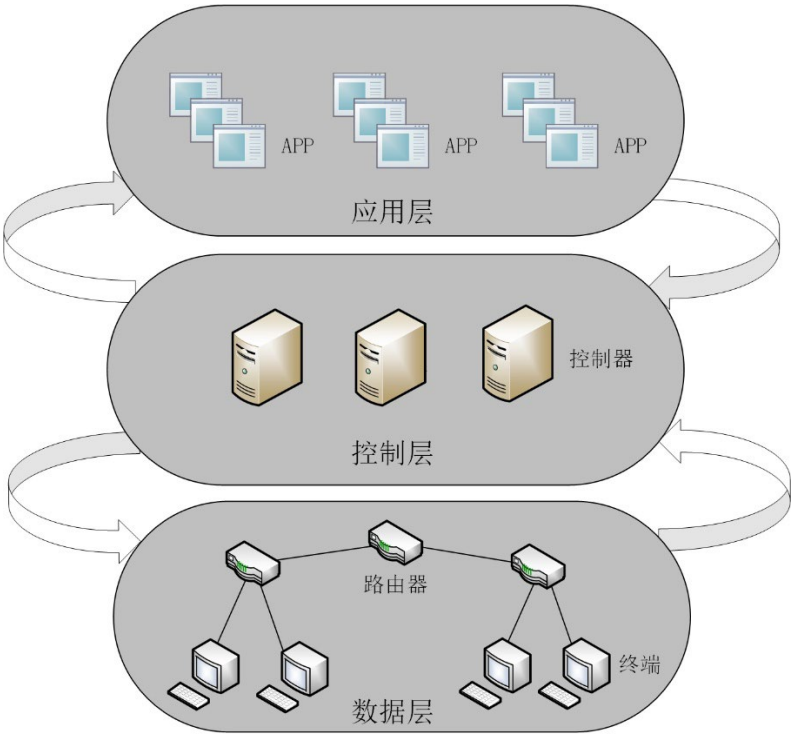

图 2-1 SDN 基本架构

SDN 网络在相当程度上考虑到如何应对未来网络架构的发展趋势，即运用网络控制权的集中化思想[36]。集权化的控制方式不仅有助于实时获取网络的全局

视图，还通过控制器与转发设备的职能分离，增强了网络资源调配的灵活性。同时，控制器对分布式转发设备的统一管理使得网络管理趋于标准化和高效化。在 SDN 网络架构中，控制平面作为核心组成部分，通常由一个或多个集中式控制器构成，其主要职责是对整个网络的运行状态进行集中管理与策略控制。控制器通过与网络中的交换机等数据平面设备进行通信，向其下发控制指令，从而实现对网络行为的统一调度与功能配置。与传统网络中分布式控制方式相比，SDN 所采用的集中式控制机制显著简化了网络配置流程，提升了故障诊断与性能优化的效率。在该架构下，数据平面则由具体承担数据包处理任务的网络设备组成，如交换机、路由器等。其在接收控制器下发的转发策略后，执行相应的数据包转发与处理操作，实现对网络流量的动态调控与精细化管理。在控制器与数据平面之间，OpenFlow 协议被广泛采用作为南向接口的标准协议之一，其定义了双方通信的数据格式与交互流程。借助 OpenFlow 协议，控制器能够灵活地向交换机下发流表项，制定精确的转发策略，从而实现对网络行为的可控性与灵活性。

SDN 网络架构在提供灵活网络控制能力的基础上，它也为网络虚拟化技术的融合与应用奠定了技术基础。其借助网络虚拟化机制，底层物理资源可被抽象映射为多个逻辑上相互隔离的虚拟网络，实现资源的细粒度划分与按需动态调度，从而提升了整体资源利用率，并强化了多租户环境中的安全隔离能力。同时，SDN 通过开放可编程的控制接口，使网络管理者能够依据具体业务需求灵活定义网络行为与调度策略，增强了系统对异构应用场景的适应能力与定制化支持。

2.2 DDoS 攻击技术

DDoS 攻击的核心特征在于其高度分布式的攻击方式，即通过操控大量地理位置分散的计算设备或终端协同发起攻击，显著提升了攻击的隐蔽性与防御的复杂度。在实际执行过程中，DDoS 攻击往往借助于网络系统中存在的安全漏洞，或通过恶意代码传播至大量移动终端设备，从而构建起规模庞大的僵尸网络[37]。一旦攻击指令下达，受控的僵尸主机会以高度并发的方式向目标系统发起请求，迅速耗尽其网络带宽、处理能力及存储资源，进而造成目标服务响应能力下降，甚至陷入瘫痪，严重破坏系统的可用性与稳定性。

DDoS 攻击的主要攻击方式有两种：资源耗尽型攻击和带宽耗尽型攻击。资源耗尽型 DDoS 攻击旨在大量的占用目标主机的关键性系统资源，如 CPU、内存及存储空间等，使其无法正常处理合法用户请求。当系统资源被大量占用时，目标服务器的性能会显著下降，最终可能导致服务瘫痪，无法对外提供正常服务。此类攻击常通过构造特殊数据包或利用协议漏洞进行，如 Teardrop 攻击和 Ping of Death 攻击等经典手段，均利用目标系统在处理异常数据包时的弱点进行资源

消耗。带宽耗尽型 DDoS 攻击则主要通过向目标服务器发送大规模的无效或恶意流量，迅速占满其网络带宽，使正常的业务流量无法被处理，进而导致合法用户的访问被阻断。此类攻击通常通过构建大规模的数据流或广播型流量，将目标网络链路推向拥塞状态，使网络服务不可用。在实际攻击场景中，攻击者往往会将上述两种方式结合使用，既通过耗尽系统资源降低服务器的处理能力，又通过带宽占用阻断外部通信，以提升攻击的破坏力和隐蔽性，来增加防御难度。

分布式拒绝服务攻击根据攻击方式的不同，可分为直接攻击和反射攻击[38]类型。这两种攻击模式在实施手法和攻击路径上存在显著差异，但都能有效对目标系统造成资源耗尽或服务中断的破坏性影响。此外，根据攻击流量的速率特征，DDoS 攻击可进一步划分为快速攻击和慢速攻击两种类型[39]。快速攻击通常在极短时间内以高强度的数据流量冲击目标系统，迅速耗尽其网络带宽或关键系统资源，使其瞬间陷入瘫痪状态，导致合法用户无法正常访问服务。此类攻击往往具有爆发性强、破坏力大的特点，常用于大规模瘫痪目标网络的场景。相比之下，慢速攻击则以一种更加隐蔽的方式对目标系统进行持久性消耗。攻击者通过以较低速率、长时间的方式持续占用系统资源，逐步削弱目标系统的处理能力，最终导致其无法正常提供服务。由于其攻击流量较小、难以被传统流量监测工具快速发现，这类攻击具有较强的隐蔽性和持久性。典型的慢速攻击方式包括 Slow Headers、Slow Body 和 Slow Read 等，这些攻击方法通过延迟数据传输或故意降低交互速率，持续占用目标系统的连接资源，最终造成服务不可用。两类攻击方式在实际场景中可能交替使用或结合实施，以提升攻击的复杂性和防御难度。

软件定义网络架构在为网络管理与配置带来灵活性和可编程性的同时，也为 DDoS 攻击提供了新的攻击面和多样化的攻击方式。攻击者可以针对 SDN 架构的不同层次设计特定的 DDoS 攻击策略，主要包括以下三类：

1、针对数据层的交换机攻击：攻击者可通过注入大量恶意流量或构造异常请求的方式，对网络交换机发起攻击，诱导其频繁生成流表项。当流表或缓存达到上限时，交换机将无法继续处理正常的数据流，进而引发网络拥塞或中断。这类攻击利用了数据层硬件资源有限的特点，以达到削弱网络承载能力的目的。

2、针对控制层的攻击：SDN 控制器作为网络的核心控制单元，承担着全局视角下的网络状态感知、决策制定与控制策略下发等关键职能。攻击者通常通过发送大量伪造的、不匹配的报文，诱导控制器进行频繁的计算和流表更新，从而消耗其 CPU、内存等关键资源。一旦控制器资源被耗尽，将导致整个网络的控制功能瘫痪，使数据层设备无法获得正常的指令，严重影响网络的稳定性与可靠性。

3、针对应用层的攻击：除了传统意义上的网络层攻击外，SDN 架构中的应用层同样面临潜在的安全威胁。攻击者可通过构造大量无效请求，如 HTTP 请求洪泛或 UDP 泛洪等典型 DDoS 攻击手段，直接针对应用层服务发起冲击，使正

常服务请求无法得到及时响应。这类攻击既可以单独实施，也可以与底层攻击相结合，以增加防御的复杂性。

由于 SDN 架构中控制层的集中式管理和数据层设备的简化设计，使得控制器往往成为攻击者的首选目标。一旦控制器遭到瘫痪，将对整个网络的运行造成严重威胁。因此，针对 SDN 环境下 DDoS 攻击的检测与防御技术的研究具有重要的理论意义和实际价值，对于保障 SDN 网络的安全性和稳定性至关重要。

2.3 区块链相关技术

比特币于 2008 年发布了白皮书，这一举动既意味着比特币诞生[40]，同时也将区块链技术首次呈现在大众视野中。作为比特币的底层核心技术，区块链本质上是一种特殊的数据结构，由多个封装数据信息的区块通过加密方式串联形成链式结构，从而实现维持数据的连续性和不可篡改性这两种特性。该技术采用非对称加密算法与分布式账本机制，确保了数据在存储与传输过程中的安全性、透明性和可靠性。

区块链技术的不断演进，使其在发展过程中愈发不满足于仅应用在加密货币领域。其技术逐渐扩散到其他行业。在金融行业中，区块链因其具有中心化与高可验证性的特征，能更好的处理跨境支付与供应链金融等金融应用场景；在教育行业中，该技术则由于其数据的高可靠性多用于学籍管理与学术成果的溯源验证。在物联网行业中，区块链则能够增强设备间数据交换过程中的安全性，降低信息篡改与泄露的风险。除此之外，还可将区块链与人工智能结合，使其有望在数据共享机制设计与隐私保护方面挖掘更具价值的创新方向。将多领域的技术进行融合将进一步推动产业结构优化，加速各行业的数字化转型进程。

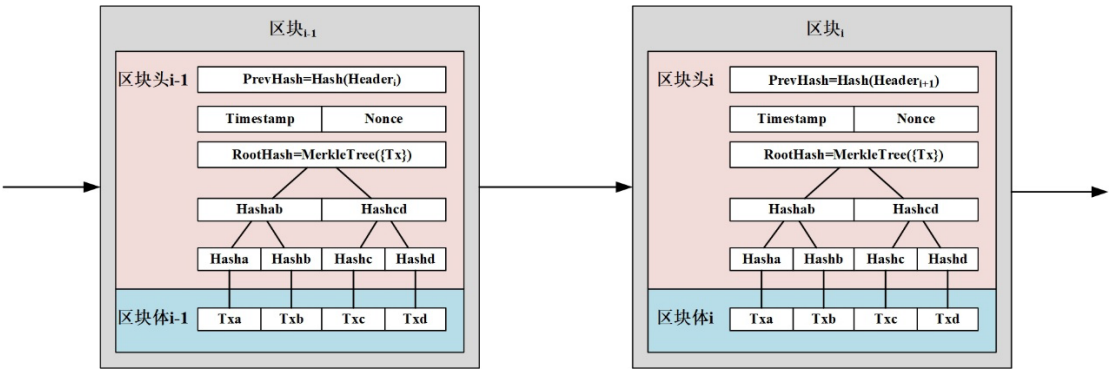

图 2-2 基本区块链原理

针对上述问题，区块链技术凭借其去中心化、不可篡改和分布式协作的特性，为 SDN 流表保护和 DDoS 攻击防御提供了可靠的解决方案。如图 2-2 所示，区块链技术允许数据以自动化、分布式的方式共享，避免了单点故障和中心化带来的安全隐患。区块链的内部结构由分层式架构组成，从下到上可依次划分为数据

层、网络层、共识层、激励层、合约层以及应用层[41]，各层之间相互协作，共同支撑区块链系统在 SDN 环境下的高效、安全与可扩展性。

数据层作为区块链的底层基础，负责存储区块、交易记录以及链式结构的关联信息，其不可篡改的链式存储结构能够有效防止攻击者恶意篡改 SDN 网络中的流表数据。同时，数据层的分布式存储结构提高了数据访问的鲁棒性，使攻击者难以通过单点攻击破坏整个网络的流表存储，提升了 SDN 网络的抗攻击能力。

网络层采用点对点（Peer-to-Peer，P2P）架构，提供了分布式节点之间的有效信息交互，避免了 SDN 网络中单一控制器成为攻击目标的风险。同时，该层集成了数字签名和哈希算法等加密机制，能有效抵御 DDoS 攻击者的流量篡改和节点伪造，保障了网络通信的稳定性与安全性。

激励层通过引入经济激励机制，促使网络中节点积极参与数据的验证与共识，增强网络的安全性与抗攻击能力。这种激励机制在 SDN 网络环境中能够激发更多节点共同维护流表数据的准确性与一致性，从而提高 DDoS 攻击下网络的整体防护水平。

合约层通过智能合约自动化地执行攻击检测与流表更新策略，智能合约的自动触发机制使 SDN 网络能够迅速响应 DDoS 攻击所产生的异常流量，减少攻击响应的延迟和人工干预带来的风险。智能合约还可保证网络管理策略和流表规则的透明性与防篡改性，从根本上强化了网络防护能力。

应用层直接负载着基于区块链 SDN 网络防御策略在现实中的运用，能够进一步结合区块链技术和无监督机器学习算法制定更有效的 DDoS 攻击实时检测方案。通过该组合方式，既可以加强攻击检测精确性，又可以有效地提高系统实时响应能力，进而实现复杂网络环境下更智能化网络安全防御。

因此，将区块链技术引入 SDN 环境中的 DDoS 攻击防御领域，能够从根本上增强网络防护能力，解决传统防御策略中心化程度高、安全性不足、实时性差等特点反映出利用区块链技术应对 SDN 网络防御 DDoS 的需求。

2.4 涉及的机器学习相关技术

随着计算机网络技术和新型通信技术的不断演进，网络流量数据呈现出高维度和大规模的特点，网络环境的复杂性也随之增加。在此背景下，由攻击行为、异常业务操作等因素引发的异常流量不断增多，给网络安全防护带来了更严峻的挑战。传统基于有标签数据的流量异常检测方法在面对大规模、高维数据时，其效果逐渐减弱。这主要是由于数据标注过程日益复杂且成本高昂，使得可用于训练的高质量标签数据越来越稀缺，进而限制了此类方法在实际场景中的适用性。为应对这一困境，近年来学术界对无监督异常流量检测方法的关注持续升温，并

在该领域开展了广泛的研究与探索。无监督方法无需依赖大量已标注的数据集，而是通过挖掘网络流量的统计特征和行为模式来识别潜在的异常情况。这类方法能够在缺乏先验知识的条件下，自主发现隐藏在复杂流量中的异常行为，因而在大数据环境下具有更强的适应性和可扩展性。

目前采用无监督算法进行异常检测的方法大致有如下 3 类：

1、基于密度法：这类方法是通过对特征空间内数据点分布密度进行分析，从而确定异常点。密度较低的数据点通常被视为异常样本，典型代表算法包括局部离群因子（Local Outlier Factor, LOF）等。

2、基于聚类方法：本方法采用聚类算法对数据进行分类，对偏离主簇群数据点进行异常判断。常用的聚类算法有 K-means、DBSCAN 等，这些算法通过衡量数据之间的相似性，从而定位异常点。

3、基于重构的方法：重构类方法通常借助自编码器等模型对数据进行降维和重构，通过分析原始数据与重构数据之间的误差来判断异常程度。若某数据的重构误差显著偏大，则可能为异常流量。

综上所述，随着网络数据规模的持续增长以及异常流量行为的复杂化，无监督异常检测方法以其良好的自适应性和无标签依赖性，已成为网络安全领域的重要研究方向之一。

基于密度的异常检测方法以数据分布特性为基础，通常假设正常样本在特征空间中呈现较高的密度分布，而异常样本则位于低密度区域。通过对样本分布密度的分析，该类方法能够有效识别偏离常规模式的数据。然而，这类方法在实际应用中面临一些挑战，尤其是在高维数据环境下，容易遭遇“维数灾难”问题，即随着特征维度的增加，数据点间的距离度量变得不再可靠，从而影响密度估计的准确性。此外，数据分布的局部密度变化也可能导致检测误差，尤其是在样本分布不均的场景下，容易出现误判或漏判。在基于密度的异常检测方法中，局部离群因子算法[42]是一种具有代表性的经典方法。LOF 通过评估样本点相对于其局部邻域的稀疏程度来判断异常性。其核心思想是使用相对密度而非绝对密度来进行异常度量，以缓解因数据密度变化导致的误判问题。具体而言，LOF 通过计算数据点与其 K 近邻的距离，评估样本在局部区域内的离群程度。当某个样本点的局部密度显著低于其邻域样本的平均密度时，该点被判定为异常样本。

基于聚类的异常检测方法通过分析数据在特征空间中的分布模式，以识别偏离正常聚类结构的异常样本。该类方法通常假设：正常样本倾向于聚集形成密集的簇群，而异常样本则不隶属于任何聚类，且远离主要的簇中心。基于这一假设，异常点可以通过其与聚类中心的距离或其在聚类结构中的孤立性来判定。在基于聚类算法的检测方法中，对于基于密度的空间聚类算法（Density-Based Spatial Clustering of Applications with Noise，DBSCAN）[43]，这种算法是一种能够被广泛

应用的经典算法。DBSCAN 的核心思想是基于数据的局部密度进行聚类，其能够识别出任意形状的稠密数据簇，同时将密度较低的点视为噪声点。DBSCAN 在处理异常检测问题时，往往将被判定为噪声的样本视为潜在的异常点。该方法具有良好的鲁棒性，尤其适用于存在不规则形状聚类的数据集。

基于重构的异常检测方法通过分析数据样本在低维空间中的重构能力，以识别潜在的异常点。该类方法的核心假设是：正常样本具有较强的可重构性，在经过降维和再构建的过程中能够保持较小的误差；而异常样本由于分布稀疏或结构特殊，往往难以被精确重构，因此其重构误差较大。这一特点可被用于异常检测，通过计算样本的重构误差来评估其异常程度。常用的基于重构的异常检测方法主要包括例如主成分分析（Principal Component Analysis, PCA）[44]算法、自编码器（Autoencoder, AE）[45]算法及其衍生的模型[46],[47]。

Liu 等人[48]，提出的孤立森林算法（Isolation Forest, IF）为异常检测领域提供了一种与传统基于距离或密度方法截然不同的解决方案。与依赖相似性度量的算法不同，孤立森林算法通过孤立样本点的方式来识别异常数据。这一方法专为异常检测设计，具备少量参数依赖和无监督学习的特点，使其在大规模数据场景下表现出色。孤立森林算法的核心由以下假设导出：异常样本在特征空间中往往具有稀疏性和与众不同的特征，因此在数据划分过程中更容易被隔离。算法通过随机选择特征和随机设定划分阈值，递归构建随机二叉树（Isolation Tree, iTree），每个样本被“孤立”所需的划分次数即为其路径长度。正常样本由于位于数据的密集区域，需要更多的划分才能被孤立，而异常样本由于与其他数据点的偏离性，通常会在较少的划分步骤中被隔离。最终，孤立森林通过计算样本在多棵随机树上的平均路径长度，以此来衡量其异常程度。孤立森林算法通常在以下四个方面具有优势：

1、线性时间复杂度：与数据量成线性关系的时间复杂度（$O(nlogn)$）使得该算法可扩展性强，适用于大规模数据集。

2、低内存消耗：由于每棵随机树仅依赖少量样本即可构建，算法在内存使用上具有显著优势。

3、少量参数依赖：算法只需设定树的数量和每棵树的子样本数，使模型调优过程简化。

4、适用于高维数据：孤立森林在高维空间下依然能保持较高的检测性能，能有效发现离群点。

在网络流量异常检测领域，孤立森林算法尤其适用于数据分布不均衡的场景，如网络流量中异常数据往往仅占少数，这种稀疏性正符合孤立森林的异常检测特性。然而，在实际应用中，数据标签往往存在噪声干扰，过多的无效或错误标签可能对分析结果造成负面影响。因此，在使用孤立森林进行异常检测之前，有必

要对数据集进行预处理，通过数据标签筛选等方式降低噪声水平，以提升模型的检测精度。鉴于孤立森林在高维数据环境下的优秀表现，本文提出的策略中就使用了孤立森林算法来进行异常流表检测。该策略基于 SDN（软件定义网络）环境中的流表数据特征，通过使用复核特征筛选算法，对数据进行预处理，减少数据标签过多造成的噪音影响，利用孤立森林高效识别异常流量，以应对网络中日益复杂的 DDoS 攻击等安全威胁。通过构建多棵随机隔离树对流表数据进行划分，该方法能够快速发现潜在异常流量，并在保证检测精度的同时有效降低计算开销，从而提升了 SDN 环境下整体的网络安全防御能力。

2.5 本章小结

本章主要介绍了与 SDN、DDoS 攻击、区块链及相关机器学习技术相关的理论基础。首先，阐述了 SDN 架构的基本原理及其核心特性，包括转发与控制分离、集中式控制和支持可编程性，这些特性使得 SDN 在网络管理和创新方面具有显著优势。其次，分析了 DDoS 攻击的主要类型及其对 SDN 架构的潜在威胁，重点讨论了针对 SDN 各层的攻击方式。接着，介绍了区块链技术的核心概念及其在各行业的广泛应用，尤其是去中心化的账本机制如何提高数据的安全性和透明性。最后，本章还探讨了无监督异常流量检测技术，特别是孤立森林算法在网络安全中的应用，展示了其在 SDN 环境下检测 DDoS 攻击和异常流量的潜力。

3 基于区块链的 SDN 流表保护研究

3.1 问题描述

在软件定义网络架构中，流表存储是网络控制与数据转发功能的核心组成部分。然而，传统的 SDN 流表存储通常依赖于中心化的控制器，这种方式虽然实现简单，却存在单点故障、数据易被篡改和安全性不足等问题，尤其在面对分布式拒绝服务攻击时更为脆弱。攻击者可通过发送大量恶意流量填满交换机流表缓存，或直接消耗控制器资源，导致网络瘫痪，严重影响 SDN 网络的稳定性和安全性。当 DDoS 攻击大规模涌入时，传统存储方式难以有效应对高流量和高并发攻击，流表存储系统往往成为攻击的薄弱环节。

为有效地应对以上挑战，区块链分布式账本特性为解决这些问题提供了切实可行的方案。利用区块链技术去中心化，防篡改以及可追溯性等特性，能够有效地加强流表数据安全，避免数据篡改以及恶意访问，同时也提高了跨域流表的共享效率，较好地解决了 DDoS 攻击中流表存储和访问延迟等问题。采用区块链技术后, SDN 控制器可以不依赖于单一控制点安全存储与共享流表数据，减少了单点故障造成的安全危险，增强了网络整体防御能力。

针对传统 SDN 流表存储的弊端，本文提出了一种使用区块链技术了进行流表存储架构。该架构设计了管理 SDN 流表的智能合约，可对流表在区块链中的读取进行高效处理。该架构特征有如下三点：

1、流表存储安全性：SDN 架构使用控制器来处理流表，但缺乏安全性考量，容易被外界篡改流表来影响 SDN 网络的正常运行。通过本文提出的设想，将流表数据存储于区块链中，借助区块链不可篡改的特性与共识机制，就可以有效提升 SDN 流表的安全性。

2、高效访问流表：区块链虽然提供安全，但是读写效率低下。为了克服高并发环境中的访问限制，我们设计了一种基于智能合约的流表存储和访问策略，确保控制器在频繁读取流表时能够保持良好的性能。

3、域间数据共享：传统 SDN 架构下，控制器间的流表数据难以高效共享。通过区块链的分布式账本特性，各控制器可在无需第三方的情况下实现安全、高效的流表数据共享，从而提升跨域流量管理能力。

本章将围绕以上三个问题展开具体研究，首先详细介绍该设计的系统模型，进而提出相应的验证体系，讲解其中的算法设计并展示依据此模型构建的实验评估结果，以充分说明区块链技术如何在 SDN 流表存储中提供安全性保障，同时兼顾高效的访问性能。

3.2 系统模型

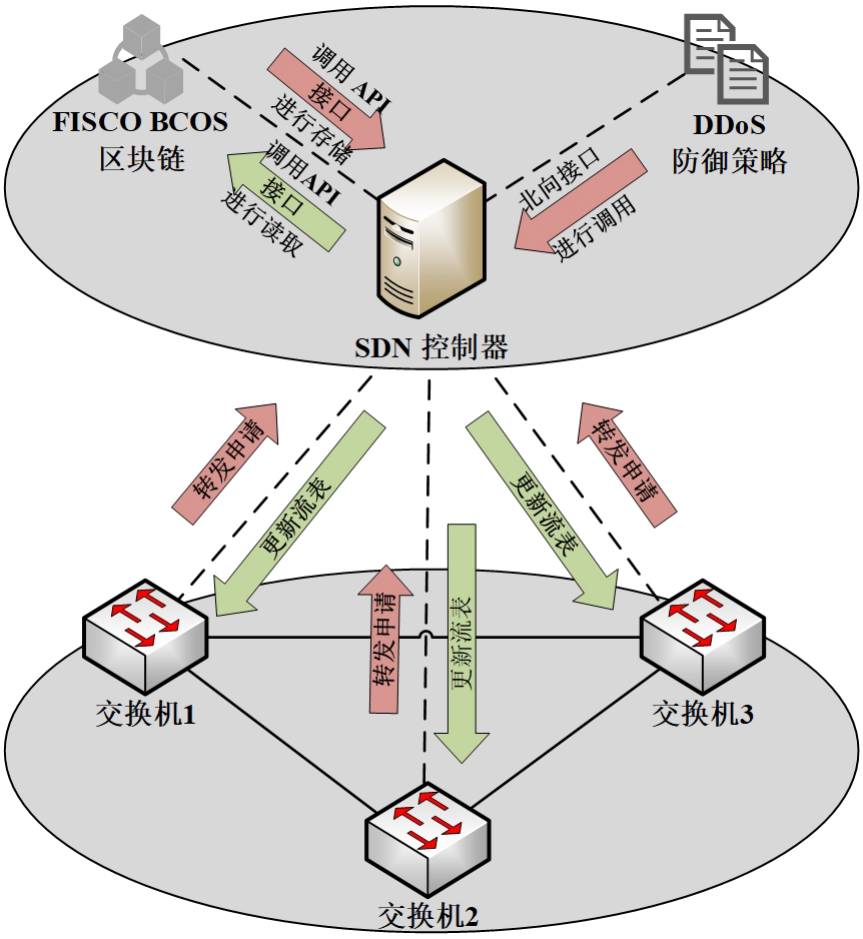

图 3-1 SDN 部署防御策略架构图

为了解决传统 SDN 流表保护中存在的单点故障、数据篡改以及跨域数据共享的难题，本研究提出了一种基于区块链的 SDN 流表保护模型。该模型以 FISCO BCOS 联盟链为底层区块链平台，通过 SDN 控制器与区块链的深度集成，构建了一个安全、透明且高效的流表管理架构，从而提高了 SDN 网络的稳定性和安全性。图 3-1 展示了该系统模型的整体架构。系统架构主要由以下核心组件构成：

1、SDN 控制器：作为网络控制和流量调度的核心，SDN 控制器通过北向接口与区块链平台和 DDoS 防御模块进行交互，并通过南向接口与数据层的交换机以及下辖的各个移动终端进行通信。控制器不仅负责流表的下发和更新，还承担着 DDoS 攻击监测与防御的任务。

2、FISCO BCOS 区块链平台：作为流表数据的分布式存储容器，区块链平台通过 API 接口与 SDN 控制器交互，负责流表数据的安全存储与权限管理。每条流表数据在被存储前需经过智能合约的验证，并在链上形成可追溯的记录，确保数据的完整性和安全性。

3、DDoS 防御模块：该模块与 SDN 控制器的北向接口相连，实时监测网络

流量状态。当检测到异常流量或潜在的 DDoS 攻击时，防御模块将触发预设的防御策略，如流量清洗、限流等操作，并通过控制器对网络拓扑进行相应调整。

4、底层交换机：作为数据平面的核心组件，交换机根据 SDN 控制器下发的流表进行数据包的转发。当交换机遇到未知流量或需要更改流表时，会向控制器发起转发申请，并在获得新的流表规则后进行更新。

这种通过区块链 API 的交互允许 SDN 控制器安全地存储和管理流表。每个 SDN 控制器的拓扑被表示为区块链，其中区块链中的每个节点对应于拓扑中的一台交换机。在区块链上部署智能合约来管理流量表，确保流量规则的安全透明存储。这些智能合同执行多项功能，包括授权 SDN 控制器、更新本地流表信息以及验证参与者访问权限。区块链的去中心化性质为流表数据提供了安全的环境，防止未经授权的访问和篡改，从而增强了 SDN 网络整体安全性和可靠性。

3.3 算法设计

本节将详细概述我们使用区块链存储 SDN 流表的设计，并通过区块链和 DDoS 检测算法处理流表数据。在 SDN 的北向端口上，我们部署了 Fisco BCOS 联盟区块链，用于为流表数据提供安全的存储环境。为了提高流表数据的处理效率，我们还引入了基于智能合约的处理机制，使得流表数据的管理变得更加高效、透明和不可篡改。域间的智能合约用于周期性间隔内收集有关数据的相关信息，并确保数据的完整性和安全性。

在 SDN 网络的典型架构中，数据平面通过南向接口向 SDN 控制器发送转发数据请求，SDN 控制器需要访问存储在流表中的数据才能准确控制数据平面完成转发操作。与传统的 SDN 控制器流表存储方式不同，我们的设计使用了 FISCO BCOS 联盟链[49]，并将智能合约[50]部署到区块链中。这样，SDN 控制器在访问流表数据时，能够通过区块链的透明机制实现数据存储与访问的高效性和安全性。在我们的设计中，SDN 控制器通过区块链 API 与 FISCO BCOS 客户端进行交互，存储流表信息，并在每次数据转发时根据当前流表规则进行数据传输。使用区块链技术实现流表存储的关键优势在于，其去中心化的特性保证了流表数据的不可篡改性与透明性，避免了传统存储方式中可能存在的单点故障问题。智能合约应满足以下几个基本功能：

1、智能合约的所有者需要拥有确定参与者是否被授权访问区块链的能力。通过智能合约，SDN 控制器可以确认其对流表的访问权限，并依据授权结果执行相应的数据存储或更新操作。

2、经过授权的参与者可以更新本地流表信息。智能合约确保只有经授权的 SDN 控制器能够在区块链上更新流表数据，而未经授权的控制器则无法执行修

改操作。

本文提出了一种特别的智能合约设计。该智能合约使用 Solidity 编程语言实现，具有灵活的数据存储和访问机制。以下是智能合约的关键方法伪代码：

协作者判定函数通过输入外部参与者用户 user 参数，来判断外部账户是否是已授权的用户。若参与者的 user 账户存在于智能合约中，则返回 true，表示该用户已得到授权；否则，将会返回 false。这一机制确保了只有授权的控制器能够参与流表数据的管理。在区块链中，数据以字符串的格式存储，每个存储条目表示为一个字符串，智能合约通过一个数组来管理所有的存储值。初始状态下，每个存储条目的值为“0,0”，表示初始状态，可见算法 1。

Algorithm 1 协作者判定

Input: 用户地址（user）
Output: 是否是协作者的判断结果，返回布尔值

```
if 协作者地址列表 (caddr) 为空 then
    return false
end if
索引 (index) ← 协作者映射表 (co[user]).索引 (index)
if 协作者地址列表 (caddr)[索引 (index)] == 用户地址 (user) then
    return true
else
    return false
end if
```

获取键或值函数最重要的作用就是为数据检索提供了一种有效而灵活的方法，它不仅能通过索引从存储数组上直接得到数据条目，还可通过对存储数据进行匹配来找到其相应索引位置，保证区块链存储数据可访问性及高效管理。这种方法只接受一个输入参数，这个参数可以是无符号整数（index）或者字符串（value），其中索引的作用是直接确定数据的存储位置，并利用字符串值对整个存储数组匹配搜索找到对应数据索引。方法实施时先对输入参数进行类型检查来决定检索方式。当输入的参数是一个无符号整数（index）时，该方法会检验所提供的索引是否处于存储数组 values 的有效界限之内。若索引成立，则该方法直接返回 values[index]以保证可以精确地得到所存储数据信息。若索引在数组范围之外，该方法会返回表示索引失效的错误信息，以避免对未定义的数据区域进行非法存取或者读取。当输入的参数是字符串（value）时，该方法会遍历存储数组 values,并逐个检查每个存储值的前缀部分是否与输入值相匹配。系统在此过程中计算出一个固定的输入字符串长度，对于 values 数组的每一个存储值截取对应长度的一段前缀，再利用哈希比较来实现匹配。在查找到匹配项时，返回存储值索引位置以指示输入值是否在存储数组内并且提供它们准确的存储位置。若遍历所有存储数组之后仍然没有找到匹配值时，该方法会返回一个错误消息表示该值没有，这样就避免了返回错误或者数据结果不确定。

所设计的方法在保证数据准确检索的前提下，增强区块链存储数据可访问性，提高管理效率。通过支持索引查询与值匹配查找相结合的方式，使智能合约能够在各种应用场景中灵活运用这一功能。比如在以区块链为基础的去中心化存储系统里，用户能够通过索引有效地获取已知数据，还能通过给出特定值找到自己的存储位置以增加数据定位时的便捷性。另外，本方法在存储和查询过程中保证安全，避免越界访问或者非法数据读取，结合区块链智能合约不可篡改的特点进一步提高数据管理可靠性。在实践中，这种方法可以适应各种区块链存储的需要，保证了用户对存储数据的高效准确的获取。它的具体算法逻辑可见于算法 2,详细阐述了以索引为基础、以值为单位的查询流程。

Algorithm 2 获取键或值

Input: 查询类型（queryType），键（key）或值（value）
Output: 查询的结果，返回对应的值或键，若不存在则返回错误信息

if 查询类型 (queryType) == ”键查值” **then**
    if 键 (key) < 值列表 (values)¨($length$) **then**
        结果 (result) ← 值列表 (values)[键 (key)]
    else
        结果 (result) ← “键不存在”
    end if查询类型 (queryType) == ”值查键”
    固定长度 (fixedLength) ← 值 (value) 的字节长度 (length)
    结果 (result) ← “值未找到”
    for i ← 0 **to** 值列表 (values) 长度 (length) **do**
        存储值 (storedValue) ← 值列表 (values)[i]
        存储值字节 (storedValueBytes) ← 存储值 (storedValue) 转换为字节
        if 存储值字节 (storedValueBytes) 长度 (length) < 固定长度 **then**
            continue
        end if
        前缀 (prefix) ← 存储值 (storedValue) 的前固定长度 (fixedLength) 个字节 (bytes)
        if hash(前缀($prefix$)) == hash(值($value$)) **then**
            结果 (result) ← i
            break
        end if
    end for
else
    结果 (result) ← “无效的查询类型”
end if
return 结果 (result)

设置或更新字符串函数的主要功能是动态管理区块链存储数组中的数据条目，既可以对现有条目进行更新，也可以在数据数组中添加新的存储值，从而实现数据的灵活管理与存储优化。该方法接受两个输入参数：无符号整数索引（index）和字符串新值（newValue），其中索引用于指示需要更新或插入的存储位置，而新值代表要存储或修改的数据内容。在函数调用过程中，系统首先对输

入参数进行校验，确保索引的有效性以及新值的正确性，以防止无效操作或数据异常影响区块链存储的稳定性。可见算法 3。

索引位于存储数组有效区间内时，本方法对索引上的数据进行直接更新得到新值以保证原数据可以进行精确替换，进而保证区块链存储完整性以及数据一致性。这一更新机制特别适合于需要更改现有数据的场合，例如更改用户的个人信息、交易历史或智能合约的参数。为进一步提升存储管理效率，本方法首先对更新后的数据和当前存储值进行核对，如果新值和原值保持一致则不需要进行更新操作，由此降低了不必要的存储开销并优化了智能合约执行开销。若所提供索引超过当前存储数组长度，本方法认为索引是新数据加入请求。该方法将新字符串值自动加入到存储数组结束，从而扩大存储空间并动态地增加数据。这种方式的灵活性体现在既能满足数据修改又能实现数据扩展，从而使得区块链存储管理变得更有效更方便。实现时，本方法从索引检查到数据赋值再到存储结构扩展几个环节来确保存储过程稳定可靠，也从数据管理层面减少维护成本，避免传统存储方案可能产生的冗余数据，使区块链存储机制更有效率。

Algorithm 3 设置或更新字符串

Input: 操作类型（operationType），索引（index），新值（newValue）
Output: 操作结果，返回是否成功

```
if 操作类型 (operationType) == ”设置” then
    值列表 (values) 追加 (append) 新值 (newValue)
    结果 (result) ← “设置成功” 操作类型 (operationType) == ”更新”
    if 索引 (index) ≥ 值列表 (values)¨(length) then
        结果 (result) ← “索引超出范围”
    else
        值列表 (values)[索引 (index)] ← 新值 (newValue)
        结果 (result) ← “更新成功”
    end if
else
    结果 (result) ← “无效的操作类型”
end if
return 结果 (result)
```

删除或清空所有字符串函数旨在提供灵活的数据管理方式，允许用户删除特定数据项或清空整个存储数组，来确保流表存储结构具有有效性和可复用性。该方法接受两个输入参数：一个无符号整数 索引（index），用于指定要删除的数据条目的位置；一个布尔值 清除标志（clearAll），用于指示是否应清空整个存储数组。当调用此方法时，首先检查 clearAll 参数的值，如果其值为 true，方法将忽略 index 参数，直接清空存储数组 values，然后重新初始化该数组，并插入默认值"0,0"。这样做的目的是确保存储数组始终保持有效的结构，避免因清空导致的数组空值异常，同时便于系统后续存储新的数据条目。如果 clearAll 为 false，方法会验证提供的 index 是否在 values 数组的有效范围内。如果索引超出范围，则

方法将返回错误信息，防止非法访问。如果索引有效，该方法会用数组的最后一个元素替换指定索引位置的数据项，并减少数组的长度，从而有效地删除该索引处的数据。这种操作方式确保删除操作的时间复杂度保持在 O(1)，提升智能合约的执行效率，并节省计算成本。综上，该方法通过两种删除模式（按索引删除与整体清空），为存储数组提供了一种高效、可控的管理方式，同时也确保了数据的完整性和可操作性，并对区块链存储资源进行了优化。

除此之外，该系统还触发了一个用于记录删除行为的删除事件。如果索引没有位于范围之内，则会返回错误的信息，说明索引已经超过边界。通过此联合方式，本功能保证对区块链上存储的数据条目进行灵活高效的管理，保持数据完整性，支持目标删除与彻底重新定位。这个功能的算法可以从算法 4 里找出来。

Algorithm 4 删除或清空所有字符串

Input: 索引（index），清除标志（clearAll）
Output: 操作结果，返回是否成功

if 清除标志 (clearAll) == true **then**
    删除存储数组 (values) 的所有条目
    存储数组 (values) 重新初始化，并插入默认值 ”0,0”
    结果 (result) ← “存储数组已清空”
else
    if 索引 (index) ≥ 存储数组 (values)¨(*length*) **then**
        结果 (result) ← “删除失败：索引超出范围”
    else
        存储数组 (values)[索引 (index)] ← 存储数组 (values) 的最后一个元素
        存储数组 (values) 长度 (length) ← 存储数组 (values) 长度 - 1
        结果 (result) ← “删除成功”
    end if
end if
return 结果 (result)

方法中伪代码可以实现智能合约在 SDN 控制器流表中进行数据存储，更新和删除的基本功能。该机制实现了区块链对 SDN 网络中流表数据透明性和安全性的保障以及仅有授权控制器可以执行流表修改动作。

3.4 实验评估

本章提出的一种基于区块链的 SDN 流表保护设计，我将从创建结合区块链和 SDN 技术搭建仿真环境和网络拓扑，转发网络流稳定性，该拓扑的可扩展性，来进行验证。本章节主要介绍实验验证体系，相仿真环境和网络拓扑、流表数据转发与原始控制器的比较、SDN 网络拓扑的可扩展性测试三部分。

3.4.1 验证体系

在这一小节中，对该策略的验证和评估分为部署稳定性和转发流表性能来进行验证。接下来将对这两点分别进行定义。

$$\mathrm{MDev} = \frac{1}{N}\sum_{i=1}^{N}\left|RTT_i - \overline{RTT}\right| \qquad (3-1)$$

策略部署稳定性：确保策略部署的稳定性对于防止干扰 SDN 网络的正常运行至关重要。我们使用数据包往返时间（Round-Trip Time，RTT）的平均偏差（Mean Deviation，MDev）来衡量这种稳定性。较低的偏差表明较高的稳定性。MDev 值表示了网络中的数据包往返时间的平均偏差数值。该值的计算过程包括收集 RTT 样本 RTT_i，计算它们的平均值 $\overline{\mathrm{RTT}}$，确定$\left|RTT_i - \overline{RTT}\right|$每个样本相对于平均值的绝对偏差，然后计算所有绝对偏差的平均值以得到 MDev。此指标的计算公式如公式（3-1）所示。

$$T_{\text{ping\_all}} = \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} T_{\text{ping}}(h_i, h_j) \qquad (3-2)$$

在可扩展性方面，我们也对策略部署的稳定性进行了检验。我们依次增加 SDN 网络拓扑中的交换机数量，以及每个交换机连接的主机数量，并记录了每次拓扑变化后，使用我们策略的 SDN 控制器对调控拓普中数据传输的时间，这里用一次 ping all 操作所需的时间来作为衡量指标，时间越短，表示使用发策略后的网络拓扑可扩展性越好。其中，$T_{\text{ping_all}}$表示所有主机互相 ping 通所需的平均时间，N为主机总数，$T_{\text{ping}}(h_i, h_j)$表示主机$h_i$成功 ping 通主机$h_j$所用的时间。该指标反映了在我们策略下 SDN 控制器调节数据传输的效率，其计算方程如公式（3-2）所示。

本小节以 Pingall 延迟与 Mdev 均值作为量化指标，从实验上分析了区块链引入对 SDN 流表访问延迟与稳定性的影响，体现了引入 FISCO BCOS 后的性能变化。其指标也能根据 Pingall 延迟的长短，与 Mdev 均值的大小从侧面检验区块链的存储开销和额外延迟对流表转发的影响程度，故本章节中，采用二者作为主要实验验证体系指标。

3.4.2 仿真环境和网络拓扑

本研究的实验环境是在 Ubuntu 20.04 操作系统上进行的，采用 Mininet 作为 SDN 网络拓扑的仿真工具。Mininet[51]是一款广泛用于构建虚拟化网络拓扑的工具，能够模拟真实网络的拓扑结构，并提供网络控制器与交换机之间的交互接口。

为了控制和管理网络流量，我们选择了 Pox[52]作为 SDN 控制器。POX 控制器是一个用 Python 编写的开源控制器，广泛应用于 SDN 研究和实验中，具备良好的扩展性和灵活性。Pox 控制器能够接收来自交换机的流表请求，并根据策略下发相应的流表规则。此外，我们使用 SCAPY 工具生成模拟网络流量，模拟 SDN 网络中主机之间的通信行为。它能够生成各种类型的网络数据包，并允许我们根据实验需求自定义数据流，以便模拟不同的网络环境。

本实验中的 SDN 网络拓扑如图 3-2 所示，拓扑结构由 1 个 Fisco BCOS 区块链节点、1 个 POX 控制器、3 个 OpenVSwitch 交换机以及 6 个主机组成。每个交换机与两个主机相连接，形成三个 SDN 子网。这样设计的网络结构旨在模拟一个小规模的 SDN 环境，便于验证基于区块链的流表存储架构在 SDN 环境中的实际表现。在拓扑中，SDN 控制器不仅负责管理网络流量，还充当 Fisco BCOS 区块链的节点，承担流表数据存储的任务。我们选用了 Fisco BCOS 作为区块链平台，它是一种金融级的联盟链，具有较强的安全性、透明性和可扩展性。Fisco BCOS 通过 Python SDK 与 SDN 控制器进行交互，并使用 Webase 组件来部署智能合约。智能合约的主要作用是实现流表数据的安全存储与管理。我们利用 Solidity 语言[53]编写并编译智能合约，将其部署到区块链上，以便在 SDN 网络中实现流表的增删改查操作。

在拓扑的初始运行阶段，交换机会通过广播地址进行流表请求，而 SDN 控制器通过北向接口接收这些请求，并将流表数据存储在区块链中。交换机向控制器发送的流数据包括原始 MAC 地址等网络信息，控制器通过智能合约验证这些信息，并将其存储在 Fisco BCOS 区块链上。之后，当交换机收到数据包时，会根据已有的流表规则进行转发操作。如果交换机未能找到匹配的流表规则，它将向 SDN 控制器发出请求，获取最新的流表信息。

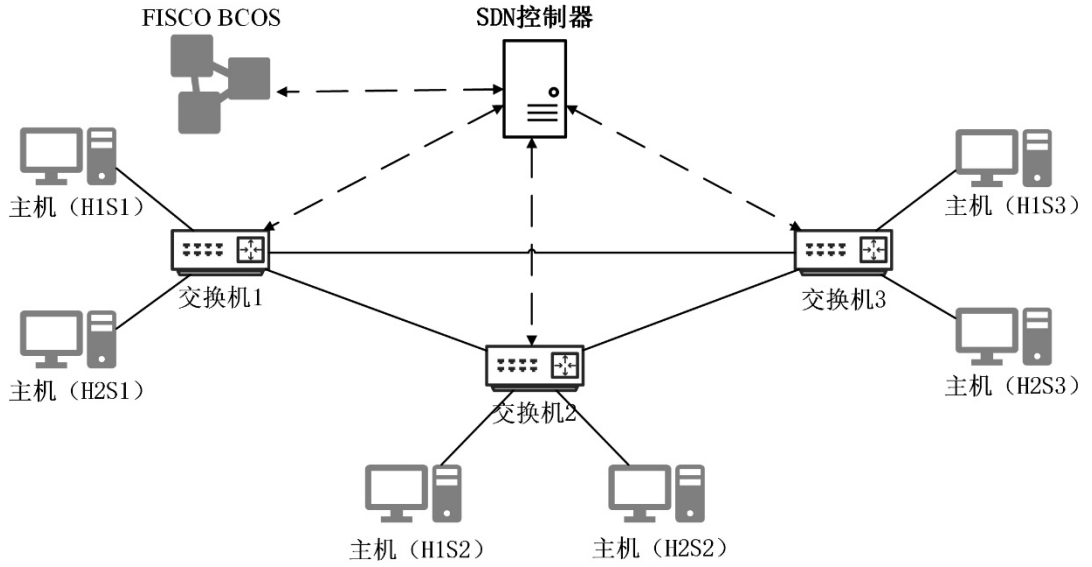

图 3-2 软件定义网络的拓扑搭建

通过这一过程，我们能够确保流表数据在区块链中的不可篡改性和透明性，同时也让各个交换机能够使用 SDN 控制器下发的最新的流表规则进行数据包转发。在区块链存储的架构下，流表数据的查询与更新都通过智能合约进行，保证了数据的安全性和准确性。在实际运行过程中，SDN 网络中的数据流成功地按照流表规则在拓扑内进行转发。通过在区块链中存储网络流表信息，我们能够实现 SDN 网络中数据流量的有效管理。区块链不仅有效保证了流表数据的安全性，还使得控制器能够及时获取流表更新，确保网络的高效运行。

本小节所创建的仿真环境和网络环境将后续将分别在 3.4.3 和 3.4.4 节中开展延迟对比实验与拓扑可扩展性分析的实验平台使用。

3.4.3 流表数据转发与原始控制器的比较

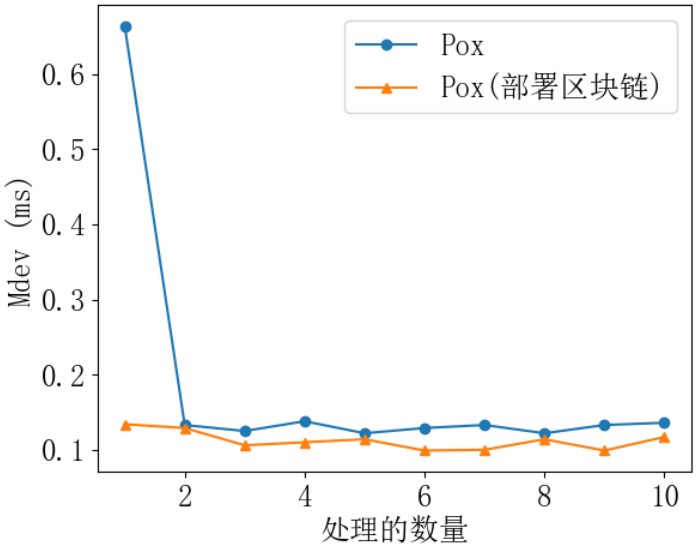

图 3-3 网络流传输稳定性的对比

在这项实验中，由于我们为 SDN 网络拓扑添加了额外的区块链架构来进行流表保护，这有可能会对网络流平台存储中控制器的正常转发流表过程造成影响，且从过往研究来看，区块链的转发效率也会因为是不同平台的区块链其内部存储数据加密过程所耗费的时间也不同。尽管我们考虑到快速转发的需求而选择了金融级国产联盟链 Fisco BCOS 来试图减轻影响，但为了评估这种影响的程度，我们将使用我们提出的设计（即使用 FISCO BCOS 区块链在 SDN 控制器北向接口搭建区块链流表存储结构后的 Pox 控制器）与未使用本策略的原 Pox 控制器进行了稳定性的比较。这里采用 Mdev 值即在主机之前传输数据时，表示往返时间的平均偏差，作为稳定性指标。试验中将两个控制器在同一拓扑环境计算拓扑中两个主机 h1s1（即交换机 s1 下的 h1 主机）与主机 h2s3（即交换机 s3 下的 h2 主机）之间 ping 连通 10 次的 Mdev 数值作为 1 组比对，共进行 10 组比对，生成的比对折线图如图 3-3 所示。

可以看出使用区块链作为流表保护容器后的 Pox 控制器在拓扑中控制数据传输的 Mdev 值基本小于原 pox 控制器的 Mdev 值，其 Mdev 的均值为 0.1122ms 低于原 Pox 的 Mdev 均值 0.1835ms。与后者相比，前者的数值降低约为 38.86%。这是因为本策略选取的区块链为金融区块链，本身适用于海量数据读取，强调交易的稳定性，此外，设计中的智能合约优化了数据处理过程，使得 SDN 控制器在访问区块链时能够保持稳定高效的操作。当 SDN 控制器从北向接口访问区块链进行流表读取时，区块链能够确保数据存储和访问的透明性与稳定性，从而提升了数据转发过程的可靠性。

综上所述，使用本文提出的 SDN 流表保护方案，不会影响到数据包的正常转发，同时在一定程度上也能提升数据包在传输过程中的稳定性，实验结果也说明，该方案在 SDN 流表保护方面具有很大的应用潜力。

3.4.4 SDN 网络拓扑的可扩展性测试

在实际应用中，SDN 网络拓扑往往会根据需求变化进行调整，拓扑的规模可能会不断扩展。为了确保基于区块链的 SDN 流表保护解决方案在不同规模的网络环境中依然能够保持高效稳定的性能，我们进行了 SDN 网络拓扑可扩展性测试。该测试的目的是评估在不同规模的 SDN 网络中，采用我们提出的区块链流表存储架构后，SDN 控制器在执行数据转发任务时是否能够保持高效的响应和稳定性，特别是在网络规模增大的情况下，流表存储与数据转发的性能是否会受到显著影响。

我们通过逐步增加 SDN 网络拓扑中交换机的数量以及每个交换机下辖的主机数量，来模拟不同规模的网络环境。具体来说，我们首先在最小规模的网络拓扑中，设置一个交换机与两个主机；然后逐步增加交换机的数量，并测试每个交换机下连接主机的数量，最终形成一个规模更大的 SDN 网络拓扑。每次拓扑调整后，我们都会执行一次ping_all操作，记录操作所需的时间，并评估控制器在不同拓扑规模下的流表转发效率。

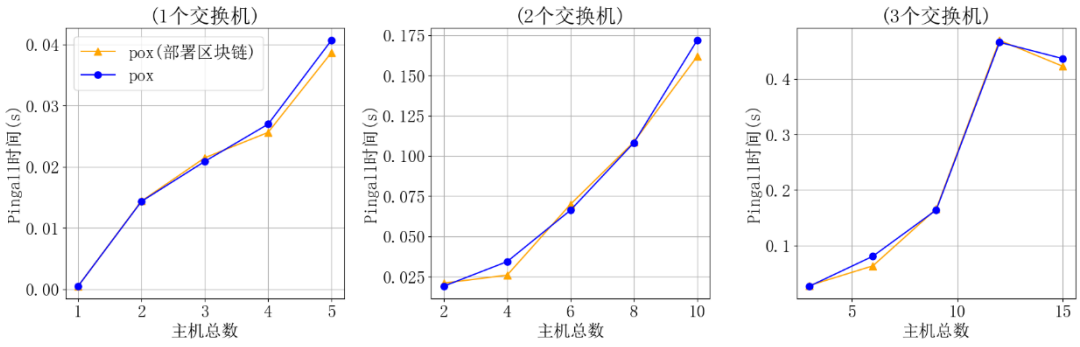

图 3-4　1 至 3 个交换机的 SDN 拓扑的 Pingall 时间比较

Mininet 仿真工具来构建不同规模的 SDN 拓扑。在每次配置调整后，使用

ping_all命令测试 SDN 控制器在拓扑变化后的响应时间。具体地，网络拓扑的规模增加包括两方面：一是交换机数量的增加，二是每个交换机下连接主机数量的增加。通过逐步增加交换机数量（从 1 个增加到 10 个），并在每个交换机上增加主机数量（从 1 个增加到 5 个），我们能够模拟出多种不同规模的网络拓扑。对于每次测试，我们记录了 SDN 控制器执行 ping_all 命令所需的平均时间。这一时间不仅反映了控制器对数据流的转发效率，还间接体现了区块链存储架构在不同网络规模下的适应性与性能稳定性。

图 3-4 所示，当交换机数从 1 个增长到 3 个过程中，随着每个交换机下辖的主机数增多，部署我们策略的 SDN 控制器执行一次ping_all时间基本和原始 Pox 控制器相差维持在 3%以内，差值较小。这表明，区块链流表存储架构在小规模网络中对控制器的性能几乎没有负面影响，控制器能够稳定高效地执行数据转发任务。

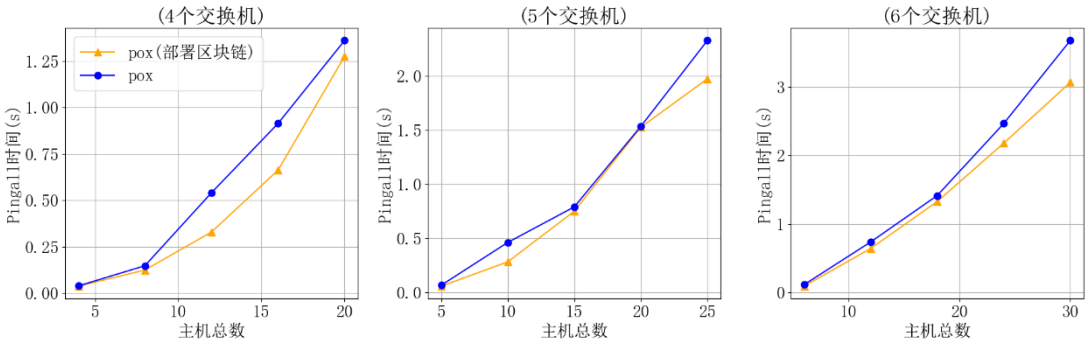

图 3-5　4 至 6 个交换机的 SDN 拓扑的 Pingall 时间比较

如图 3-5 所示，当交换机数从 4 个增长到 6 个过程中，尤其是当交换机数量从 4 个增加到 6 个时，基于区块链的流表保护架构仍然表现出较好的性能。在一些情况下，由于 FISCO BCOS 区块链的金融级优化特性，SDN 控制器的响应时间甚至优于没有区块链架构的原始 Pox 控制器。这是因为区块链在读取操作中的稳定性和高效性使得控制器在执行数据转发时，能够保持相对较低的延迟。

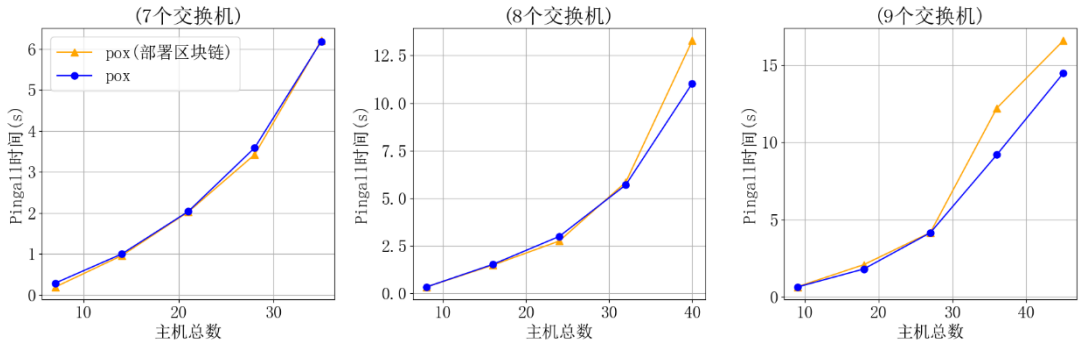

图 3-6　7 至 9 个交换机的 SDN 拓扑的 Pingall 时间比较

如图 3-6 所示网络规模进一步增加，尤其是交换机数量超过 7 个时，基于区块链的流表保护架构开始表现出一些性能瓶颈。随着交换机数量的增加，每个交

换机下连接的主机数量增多，SDN 控制器执行 ping_all 命令所需的时间逐渐增加。这是因为区块链需要处理更多的存储和验证请求，虽然通过智能合约优化了读写操作，但区块链的存储和共识机制依然不可避免地带来了某些延迟，使其在大规模的网络中性能开始出现下降。

总而言之，通过可扩展性测试，我们验证得出以下结论：在中小规模的 SDN 网络拓扑中，基于区块链的流表保护架构能够保持较高的转发效率和稳定性，甚至在一些情况下优于传统控制器。而在大规模网络拓扑中，虽然区块链存储提供了更强的数据安全性和透明性，但随着拓扑规模的进一步扩大，尤其是当 SDN 网络拓扑内交换机数量为 1 个到 7 个时，系统的访问延迟和存储瓶颈可能会对性能产生一定影响。因此，针对大规模网络的拓扑，可能需要进一步优化区块链存储架构，采用更加高效的区块链技术，或者结合其他优化手段，以确保系统在大规模环境中的稳定运行。

3.5 本章小结

本章针对传统 SDN 流表存储方式在安全性与效率方面的不足，提出基于区块链保护 SDN 流表的方法。我们对流表保护中的关键问题进行了深入探讨，特别是当面临分布式拒绝服务（DDoS）攻击时，传统的中心化控制器所展现出的弱点。区块链技术利用其去中心化、防篡改和可追溯的特性，为 SDN 流表的存储提供了一个安全且可靠的解决方案。该章研究方法将 FISCO BCOS 联盟链和智能合约相结合，实现流表数据安全存储及跨域共享。通过设计有效的智能合约及数据存储机制来保证流表数据完整性及有效访问。我们也进行了算法设计与实验评估，认为所提出的区块链流表保护架构无论从安全性还是性能上都具有优异的表现。

经过实验评估与验证，本文发现在中小规模 SDN 网络拓扑上，区块链流表保护能提供更高稳定性与数据转发性能，在一定条件下比传统控制器更好。其不足点在于，实验验证中在 7 个交换机之后系统出现明显性能瓶颈，延迟上升，这些结果表明当前方案在大规模网络环境下的适应性仍有限。考虑到现实 SDN 部署多为小中型网络，且本方案设计初衷亦是面向中小规模网络的轻量化防御场景，故本研究已实现设计初衷目标，故未进一步扩展至更大规模拓扑。未来可在后续研究中引入链分片与轻量区块链机制以提升大规模适应性。总体来看，本章节所提出的区块链 SDN 流表保护方法为传统 SDN 架构下流表的安全性，透明性及高效访问等难题提供了创新性解决方案，极具应用价值与发展潜力。

4 基于机器学习算法的双层 DDoS 检测策略研究

4.1 问题描述

在 SDN 网络环境中，分布式拒绝服务攻击通过恶意流量洪泛、流表资源耗尽或控制器通信链路阻塞等多样化手段，严重威胁到了网络可用性与网络安全性。尽管现有研究提出了基于机器学习的方法以提升攻击检测能力（如实时检测[19]、增强型 k 均值聚类[20]等），但在实际部署中仍面临两大核心问题：其一，单一检测模型难以兼顾显性攻击（如流量洪泛）与隐蔽攻击（如低速资源耗尽）的差异化特征；其二，传统特征选择方法（如皮尔逊相关系数、Lasso 正则化）因依赖单一评价指标，易导致模型泛化能力不足，尤其在面对复杂多变的攻击模式时准确率显著下降（文献[54]指出跨场景测试中性能降幅可达 18%-25%）。因此，为了提升 DDoS 检测的精度与适应性，当前亟需设计一种具备分层处理机制与多指标特征选择能力的检测策略。该策略应既具备快速识别高强度攻击行为能力，又具备能捕捉低频隐蔽威胁的关键特征，兼顾模型的实时性与计算开销控制的能力。

为此，本章提出一种基于机器学习的双层 DDoS 检测策略，通过分层次防御架构与复合特征选择机制实现攻击检测的精准性与高效性协同优化。第一层采用时间窗口算法与令牌桶机制，基于流量特征频率快速识别并抑制显性攻击；第二层结合互信息、随机森林与递归特征消除的复合特征选择算法，筛选高区分度流量特征，并利用孤立森林算法检测隐蔽攻击。

接下来，将陆续介绍该设计的系统模型、验证体系、问题建模、算法设计以及相关实验评估结果，以展示双层检测策略在提升检测精度与降低计算时间方面的双重优势。

4.2 系统模型

针对 DDoS 攻击的检测依赖于 SDN 网络拓扑中的已有网络流数据信息的存储与处理，本研究提出的整体检测架构已在上一章图 3-1 中进行了展示。在 SDN 网络中，DDoS 攻击通常通过大量恶意流量占用控制器资源，导致流表缓存耗尽、带宽受限，甚至使网络服务瘫痪。因此，为了有效识别和防御 DDoS 攻击，本研究在 SDN 控制器的北向接口部署了基于区块链的存储系统，并结合机器学习技术构建了双层 DDoS 检测策略。本策略的第一层检测采用时间序列分析和阈值计算，以实时监测并快速响应明显的攻击行为；第二层检测则使用基于机器学习的无监督算法，以识别更为隐蔽、传统方法难以检测的攻击。此外，由于 DDoS 攻

击流量通常具有高维度和复杂特征，本研究提出了一种结合互信息算法（Mutual Information, MI）、随机森林算法（Random Forest, RF）和递归特征消除（Recursive Feature Elimination, RFE）的复合特征值筛选方法，来对检测流程进行优化，从而提高检测的效率。

1、第一层策略利用时间序列分析和令牌桶算法来检测和缓解明显的 DDoS 攻击。时间序列分析用于跟踪特定时间窗口内网络流量特征的变化，主要依据源 MAC 地址、IP 地址、端口号等信息，统计这些特征在指定时间段内的出现频率，并设定阈值来识别异常流量。若某源地址的访问频率显著高于正常流量模式，即可判定其为 DDoS 攻击的潜在来源。时间序列分析的核心在于实时监测流量变化，并通过设定动态阈值确保攻击流量能够被快速发现。随后，令牌桶算法（Token Bucket Algorithm，TBA）用于限制可疑攻击源的流量速率，避免恶意流量影响正常数据传输。令牌桶机制通过预设的速率控制攻击源的数据发送，使其无法持续向控制器发送超出阈值的数据包，从而减少 DDoS 攻击对网络资源的影响。结合时间序列分析和令牌桶算法，该层检测能够在低计算开销的情况下快速发现和缓解明显的 DDoS 攻击，确保 SDN 网络的基本稳定性。

2、第二层策略针对更隐蔽的 DDoS 攻击，采用机器学习算法进行深度检测和防御。DDoS 攻击可能表现为低速、间歇性或经过伪装的异常流量，使得传统基于规则的检测方法难以准确识别。本研究提出的第二层检测策略首先通过复合特征选择方法筛选关键数据特征，以降低数据维度，提高检测效率。具体而言，本研究结合了三种特征选择算法：互信息算法用于衡量不同特征与分类目标的相关性，以确保所选特征与 DDoS 攻击的行为模式高度相关；随机森林算法通过计算各特征在决策树中的重要性，筛选出具有较强区分能力的特征；递归特征消除则通过迭代训练模型并逐步移除影响较小的特征，以确定最优的特征子集。经过特征选择后，数据维度被有效压缩，仅保留最具代表性的特征，从而减少计算开销，提高检测效率。

在完成特征选择后，系统采用无监督机器学习算法孤立森林（Isolation Forest, IF）进行 DDoS 攻击检测。孤立森林算法基于树结构构建模型，通过递归划分数据样本来识别异常点，其核心思想是异常数据比正常数据更容易被孤立，因此异常数据的路径在长度上会短与正常情况。利用孤立森林算法，可以有效检测出传统规则检测方法难以发现的隐蔽攻击流量，并对其进行标记。当流量被判定为 DDoS 攻击后，系统进一步调用令牌桶算法，对攻击源的数据传输进行限流，以减少其对 SDN 网络的干扰。

该双层检测策略可以同时兼顾 DDoS 攻击检测实时性和准确性。第 1 层检测是建立在时间序列分析与令牌桶算法基础上，能够迅速检测出明显的 DDoS 攻击，采取防御措施确保 SDN 控制器在流量过载情况下不瘫痪。第 2 层检测采用

特征选择与无监督学习算法相结合的方法实现了更加隐蔽的攻击检测，同时与流量控制策略相结合保证了网络在复杂攻击场景下仍然能够稳定运行。该系统模型设计创新点是结合区块链技术和机器学习方法，在保证流量数据安全存储的同时，增强检测策略灵活性及适应性。

4.3 问题建模

该小节阐明了本文涉及的双层 DDOS 检测策略，提供了明确的数学定义:

1、第一层防御策略包括时间窗口算法和令牌桶算法，用于检测和抑制显著的 DDoS 攻击。首先，初始数据建模涉及网络流量特征的提取。我们有一个包含网络流量信息的流表数据集，数据集中的每一行表示一个网络流记录。我们提取如下特征：源 MAC 地址(MAC_i)、时间戳（t）。

时间 2 窗口算法用于分析网络流量中特定特征的频率，在一段时间窗口内，计算每个源 MAC 地址的出现频率。定义时间窗口长度T，第i个源 MAC 地址(MAC_i)，在t时间内MAC_i出现的频率为$f(MAC_i,t)$，在时间窗口T内计算$f(MAC_i,t)$，并将频率值累计，得到时间窗口的总频率$F(MAC_i)$。其频率$f(MAC_i,t)$的计算公式可见公式（4-1）。

$$f(MAC_i,t)=\frac{N_i(t)}{N(t)} \tag{4-1}$$

其中，$N_i(t)$表示在时间t来自MAC_i的包的数量，$N(t)$表示在时间t的总包数量。将频率值累计后，得到在时间窗口的总频率可见公式（4-2）。对于任意时间t，计算时间窗口$[t-\Delta t,t]$内源 MAC 地址的出现次数$F(MAC_i)$。设置 SDN 拓扑所在网络环境的异常频率界限值θ。如果$F(MAC_i)>\theta$，则该数据包被视为 DDoS 攻击来源。

$$F(MAC_i)=\sum_{t=t_0}^{t_0+T} f(MAC_i,t) \tag{4-2}$$

令牌桶算法用于限制这些可疑攻击源生成的数据量，确保正常数据流的转发不受影响。定义当前时间t时令牌桶中的令牌数量B_t，上一时间$t-1$时令牌桶中的令牌数量B_{t-1}，令牌生成速率（每秒生成的令牌数量）R，当前时间与上次更新令牌桶数量之间的时间间隔Δt，某攻击源MAC_i需要转发的数据包数量D_i。令牌桶中的令牌数量随着时间的推移而增加，如公式（4-3）所示。

$$B_t=B_{t-1}+R\times\Delta t \tag{4-3}$$

若某攻击源MAC_i需要转发的数据包数量D_i超过当前令牌桶中的令牌数量B_t，

则丢弃部分数据包，使其转发的数据量不超过令牌桶中的令牌数量$D_i \le B_t$。对于任意时间 t，计算时间窗口$[t-\Delta t, t]$内某特征值（如源 MAC 地址）的出现次数n。如果$n > \theta$，则该数据包被视为 DDoS 攻击源，并采用令牌桶算法进行抑制。

2、第二层防御策略结合复合特征选择算法和隔离森林算法，用于检测不易识别的 DDoS 攻击。首先，复合特征选择算法预处理存储在区块链中的网络流信息，确定数据特征值的最佳数量和检测顺序。特征集合$F = \{f_1, f_2, \ldots, f_n\}$，其中$f_i$表示第$i$个特征。复合特征选择算法计算特征重要性得分可见公式（4-4）。

$$X_{\text{score}} = \sum_{i=1}^{n} w_i \cdot S_i(f) \tag{4-4}$$

其中，w_i表示第i个特征选择方法的权重，$S_i(f)$表示特征选择方法i对特征f的重要性评分。为了使w_i的选取避免过度主观，于是采用归一化过程进行调整。归一化特征值系数计算可见公式（4-5）。$S_{\max}$和$S_{\min}$分别是特征选择方法得分的最大值和最小值。

$$w_i = \frac{S_i(f) - S_{\min}}{S_{\max} - S_{\min}} \tag{4-5}$$

最终经过归一化处理的复合特征选择算法的计算特征重要性得分的公式可见公式（4-6）。

$$X_{\text{score}} = \sum_{i=1}^{n} \frac{S_i(f) - S_{\min}}{S_{\max} - S_{\min}} \cdot S_i(f) \tag{4-6}$$

根据X_{score}进行排序，选择前k个特征$\{f_1, f_2, \ldots, f_k\}$。接着使用无监督学习孤立森林算法来计算得出异常值。其基本思想是异常值更容易被孤立。隔离森林算法通过构建多个随机树（Isolation Trees），将数据逐渐划分，从而计算每个样本的孤立评分。数据集$D = \{d_1, d_2, \ldots, d_m$，其中$d_i$表示第$i$条数据；树的数量$t$；样本$d_i$的平均路径长度$h(d_i)$；样本$d_i$的孤立评分$s(d_i)$；孤立评分阈值$\tau$。该算法会对数据集$D$使用$t$棵树进行训练。每棵树通过随机选择特征和阈值，将数据集递归划分。然后计算每个样本d_i的平均路径长度$h(d_i)$，路径长度表示从根节点到达叶节点所需的步数。最后通过公式（4-7）得出孤立评分$s(d_i)$，其中，$c(m)$是一个归一化常数，由数据集大小m决定，计算公式见公式（4-8）。$H(i)$表示第i个调和数，见公式（4-9）.γ是欧拉常数，约为 0.577，若$s(d_i)$超过设定的阈值τ，则认为d_i是异常值。

$$s(d_i) = 2^{-\frac{h(d_i)}{c(m)}} \tag{4-7}$$

$$c(m) = 2H(m-1) - \left(\frac{2(m-1)}{m}\right) \tag{4-8}$$

$$H(i) = \ln(i) + \gamma \tag{4-9}$$

定义分类阈值θ。当孤立评分$s(d_i)$超过阈值θ时，认为样本d_i是异常值，否则认为是正常值。见公式（4-10）。

$$\widehat{y_i} = \begin{cases} 1 & if\ \ s(d_i) > \theta \\ 0 & if\ \ s(d_i) \leq \theta \end{cases} \quad (4-10)$$

分类阈值θ的确定是通过计算$F1$得分来实现的。首先，计算精准率（Precision）和召回率（Recall）随不同阈值变化的曲线。然后，使用$F1$得分的公式(4-15)，通过遍历不同的阈值，计算每个阈值对应的$F1$得分，并选择$F1$得分最高的阈值作为最终的分类阈值θ。在完成问题建模后，接下来我们详细描述具体研究方法。

4.4 算法设计

1、第一层针对容易检测的 DDoS 攻击，通过在一段时间内持续读取存储在区块链中的网络流量数据，并分析特定特征的频率，以确定是否将某一来源分类为 DDoS 攻击源。在本研究中，源 MAC 地址和存储在区块链中的时间戳被用作数据分析特征。利用时间序列算法，确定每条数据的时间戳是否落入预定义的时间区间内。读取该时间区间内的数据的源 MAC 地址，并且如果同一时间区间内特定源 MAC 地址的出现次数超过根据拓扑网络环境设定的阈值，则该来源被视为疑似 DDoS 攻击源。

随后，采用令牌桶算法对该攻击源生成的数据进行限制，通过设定桶内的令牌总数和令牌发放速率（每个令牌允许攻击源转发一条数据）来确保正常数据流不受影响。令牌桶中令牌数量的更新原则如公式(4-6)所示。令牌桶中在某一时刻的令牌数量$B(t)$每秒按照生成速率R增加，同时随着数据的转发而减少。当前令牌数量取决于前一秒的令牌数量B_{t-1}以及该时间段内新生成的令牌数量。当数据需要转发时，若当前令牌数量充足，则允许转发；否则，限制转发。

2、第二层策略针对不易检测的 DDoS 攻击。为了增强我们方法的实用性，我们使用了 CIC-DDoS2019 数据集[56]作为第二层中使用的孤立森林算法[57]的数据支持，对数据流进行了二元分类训练。该数据集包含基于 11 种不同类型 DDoS 攻击的网络流的 77 个统计特征。由于计算资源有限，且本研究采用无监督学习方法识别异常行为模式，攻击类型在建模中主要体现为特征空间中的分布差异，因此无需针对每种攻击类型分别设计检测机制，所提策略在原理上具有对多种攻击形式的天然泛化能力。在这项工作中我们仅关注与 SYN 类型攻击相关的数据，创建了一个训练集（仅包含正常流量）和一个测试集（包含正常流量和异常 SYN 流量）。为了减少使用隔离森林算法在网络拓扑中对数据流进行二元分类所需的时间，并提高检测的准确性，我们需要首先确定在使用 CIC-DDoS2019 数据集进行预训练时要选择哪些数据特征。在这里，我们提出了一种复合特征选择方法，

该方法集成了多种特征选择算法。该方法旨在通过利用不同算法的优势来增强特征选择的鲁棒性和准确性。我们使用了三种特征选择算法（见表 4-1）：互信息（MI）[58]、随机森林（RF）[59]和递归特征消除（RFE）。前两种算法用来计算每个特征的重要性分数。

表 4-1 复合特征选择策略

算法	定义
互信息	互信息衡量了特征存废对 Y 预测的信息贡献强度。
随机森林	通过构建多个决策树来降低单一决策树过拟合的风险，从而提升模型的泛化能力。
递归特征消除	通过递归训练模型并移除非重要特征，从而筛选出最优特征子集。

设$\mathrm{F} = \{\mathrm{f}_1, \mathrm{f}_2, \ldots, \mathrm{f}_\mathrm{n}\}$表示特征集，其中$f_i$是第$i$个特征。使用 MI 计算每个特征的重要性得分作为：

$$S_{\mathrm{MI}}(f_i) = I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \qquad (4-11)$$

其中X表示特征，Y表示类标签，$\mathrm{p}(X,\ Y)$是X和Y的联合概率分布。使用随机森林（RF）的每个特征的重要性分数由节点杂质的减少量确定，该减少量由到达该节点的概率加权。对于特征f_i，重要性得分计算为：

$$S_{\mathrm{RF}}(f_i) = \frac{1}{T}\sum_{t=1}^{T}\sum_{n\in N_t} I_n \cdot p(n) \qquad (4-12)$$

其中T是林中的树的数量，N_T是树T中在特征f_i上分裂的节点的集合，i_N是节点N处的杂质减少，$p(N)$是到达节点N的样本的比例。

RFE 方法通过递归地移除最不相关的特征，并应用隔离森林算法进行二元分类。之后，我们测试识别训练集中每个数据流是否为攻击源的准确性。最终，使用隔离森林算法获得最高准确率的数据特征集将被选为实时攻击识别的检测特征。该过程可以描述如下：

（1）使用互信息（MI）和随机森林（RF）计算F中每个特征f_j重要性得分$S_{\mathrm{model}}(f_i)$。

（2）在当前特征集F上训练模型。

（3）移除具有最低重要性评分的特征f_j：将特征集合更新为$F \leftarrow \{f_i\}$。

（4）重复步骤 2 到 3，直到达到所需的特征数量。

为了确保不同算法的重要性分数的可比性，我们使用最小-最大归一化来归一化分数。具体计算见公式（4-4）至（4-6）。我们通过结合互信息和随机森林算法的归一化得分，来计算特征f的综合重要性得分。我们为每个算法分配权重。设w_{MI}和w_{RF}分别为 MI 和 RF 算法的权重，则综合重要性得分$S_{\mathrm{com}}(f)$的计算公式如下：

$$S_{\text{com}}(f) = w_{\text{MI}} \cdot S_{\text{MI,norm}}(f) + w_{\text{RF}} \cdot S_{\text{RF,norm}}(f) \quad (4-13)$$

基于综合重要性得分，我们使用递归特征消除（RFE）方法递归地移除最不重要的特征，保留前k个特征，这些特征在训练后能够在测试集上使孤立森林算法的二分类准确率达到最高。所选的特征集F_{sel}由以下公式给出：

$$\boldsymbol{F}_{sl} = \{f \in F \mid S_{\text{com}}(f)\} \quad (4-14)$$

3、根据第二层检测 DDoS 策略的逻辑，该策略会将一段时间内的存储在区块链中的数据流的特征进行分析，使用孤立森林算法将数据流分为正常流量和异常流量两种类型。再调用令牌桶算法，对其异常流量的源 Mac 地址发出的数据进行限流。我们通过上文介绍的复合特征选择算法确定好了具体需要孤立森林算法分析的一条流表数据的哪些特征，当 SDN 网络拓扑运行时，为了尽可能减少判断流表的过程影响到正常流表转发效率，SDN 控制器会在指定的间隔时间段内访问一次存储的流表信息，并将此时存储的流表信息看作新的测试集，为每条流表数据计算出各自的异常值。

由于孤立森林算法不能直接执行二元分类，我们需要根据计算出的异常得分进行阈值划分。为此，我们在训练阶段除了提前使用了复合特征筛选算法确定了哪些流数据特征需要被孤立森林用作训练，还确定了孤立森林算法计算出的异常得分阈值，高于阈值的会被看做是正常流量，低于阈值的看做是异常流量（因为训练集只含正常流量）。该阈值具体是使用精准率-召回率曲线（Precision-Recall Curve，PRC），通过计算 F1 值作为平衡精准率和召回率的指标，来评估不同阈值下的分类效果。最后，选择使 F1 值最大的阈值作为最佳阈值，用于将异常得分划分为正常流量和异常流量。

4.5 实验评估

为了验证本章提出的一种基于机器学习的双层检测策略设计，我将从 DDoS 攻击检测的准确性，来进行验证，在此之前我也会介绍如何使用复合特征值筛选算法来寻找最佳的特征值组合进行后续可疑攻击源分析与判断。本章节主要介绍验证体系、筛选最佳特征值组合、DDoS 检测准确性验证。与同类防御策略的异常网络流判定时间比较。

4.5.1 验证体系

这一小节中，本策略将从检测准确性来设计验证体系。检测准确性是评估防御策略识别和缓解 DDoS 攻击效果的关键指标，直接影响 SDN 网络在攻击环境下的稳定性和安全性。为了确保本研究提出的双层 DDoS 检测策略能够有效检测

并应对不同类型的攻击，我们采用标准的分类评估指标，包括精确率（Precision）、召回率（Recall）和 F1 分数（F1-score），并对误报率和漏报率做相应的分析，来确保其适用性和可行性在真实环境中仍有效。

在本研究中，采用与论文[54]一致的方法，通过量化标准分类指标来评估防御策略的有效性。具体的，我们计算了这四个核心参数——真阳性（True Positive, TP）、真阴性（True Negative, TN）、假阳性（False Positive, FP）和假阴性（False Negative, FN），用来衡量该策略的表现。TP 指的是成功检测并正确识别为正常流量的数据量；TN 指的是成功检测并正确识别为异常流量的数据量；FP 是误判的情况，即正常流量被错误地标记为 DDoS 攻击流量；FN 是漏检的情况，即实际的 DDoS 攻击流量被错误地标记为正常流量。通过这四个指标，可以计算精确率、召回率和 F1 分数，从而综合衡量检测策略的分类性能。

F1 分数是精确率（Precision）与召回率（Recall）的调和平均值，其用来作为衡量分类模型准确性和完整性的指标。由于在 DDoS 检测中，精确率和召回率往往存在一定的权衡关系，因此 F1 分数能够提供一个更加全面的评价标准。如果检测策略的精确率较高但召回率较低，则说明系统可能会漏报部分 DDoS 攻击，从而降低防御效果；而如果召回率较高但精确率较低，则意味着系统误报率较高，可能会影响正常业务流量的正常传输。因此，F1 分数可以在两者之间找到一个合理的平衡点，确保策略在准确检测攻击的同时，减少误报和漏报的发生。其计算公式如下：

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{Pre} \cdot \mathrm{Rec}}{\mathrm{Pre} + \mathrm{Rec}} \qquad (4-15)$$

在（4-15）公式中的精确率Pre表示识别为正类的实例中实际为正类的比例。较高的精确率意味着较少的正常流量被误判为异常流量，以减少不必要的干扰。 它的计算公式为公式（4-16）。

$$\mathrm{Pre} = \frac{TP}{TP + FP} \qquad (4-16)$$

在（4-15）公式中的召回率Rec衡量正类实例中被正确识别为正类的比例。召回率，用于评估分类模型的全面性较高的召回率意味着较少的异常流量被漏检，从而提高防御系统的可靠性和安全性。它的计算公式为公式（4-17）。

$$\mathrm{Rec} = \frac{TP}{TP + FN} \qquad (4-17)$$

上述验证系统清楚地衡量了拟议策略的可用性。为了验证策略的有效性，本研究首先使用 CIC-DDoS2019 数据集进行实验，该数据集包含多种类型的 DDoS 攻击流量，并覆盖不同网络环境下的流量特征。将此数据集分为训练集和测试集，使用训练集来优化本文提出的检测模型，使用测试集用于评估本文策略的检测效果[55]。此外，为了模拟 SDN 网络中的实际应用场景，我们在 Mininet 仿真环境中搭建了 SDN 网络拓扑，并使用 POX 控制器管理网络流量。在实验过程中，我

们使用不同的流量负载和攻击强度，对策略的检测性能进行全面评估，并与其他 DDoS 检测方法（如基于随机森林、k-means 聚类等方法）进行对比分析，验证策略在不同网络条件下的均能有较好的适应性和鲁棒性。

总之，该研究从标准分类评估指标，误报率及漏报率多维度进行分析，并构建完善的验证体系来保证双层 DDoS 测试策略的准确性，实时性及适用性。在接下来的章节里，我们将对详尽的实验步骤和得出的结果进行深入的分析，以证实这一策略的实用性和有效性。

4.5.2 筛选最佳特征值值组合

在进行后续实验之前，我们需要对我们算法中的一些参数做好设置，以使得整个算法模型对数据集进行的二元分类能有最佳的效果。这些参数主要包含两个，一是孤立森林算法的树的数量，二是算法需要训练的数据特征。树的数量（n_estimators）这一参数用来设置模型中孤立树的个数。更多的树可以提高模型的稳定性和精确度，因为它们可以更全面地捕捉数据中的异常模式。数据特征是从网络流量中提取的有意义的信息，用于帮助模型区分正常和异常流量。选择适当的特征用来分析能够显著的提高此模型性能。我们先使用复合特征选择算法算出了数据集每个特征的相关性系数值，然后依照数值降序排列这些特征的顺序，使用孤立森林按照该顺序进行训练，并逐轮去除系数值最低的特征，每轮训练完成后使用测试集检验准确度，准确度最高对应的一组特征值和孤立森林树的数量就是我们所需的参数。

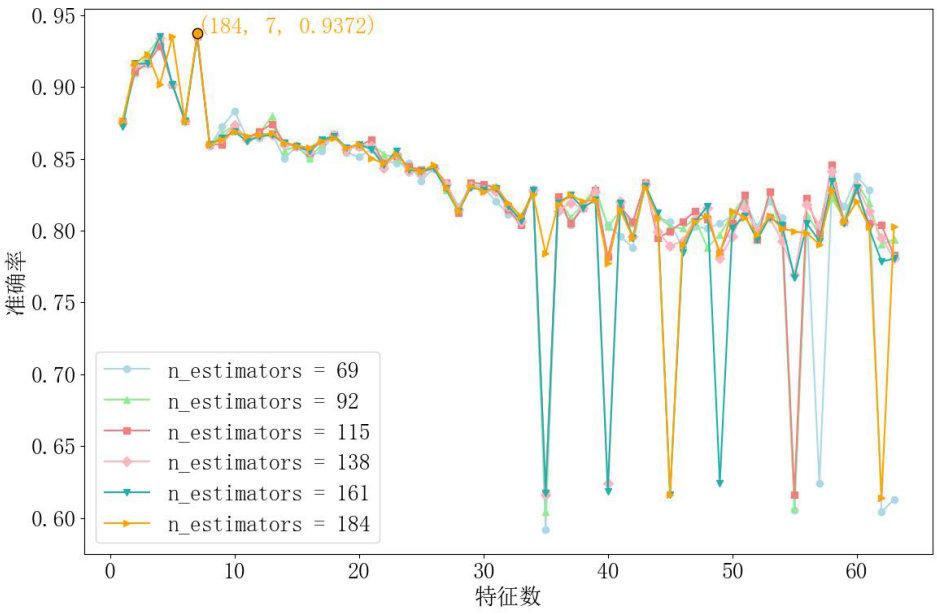

图 4-1 基于特征选择优化的算法模型参数对比分析

在识别模式和检索信息的过程中，使用真阳性（TP）、真阴性（TN）、假阳

性（FP）和假阴性（FN）的二元分类指标有助于衡量系统性能。由于 DDoS 攻击数据包识别也是一个二元分类问题，因此我们使用这些指标进行研究。在我们针对 DDoS 攻击的二元分类问题中，TP 是指使用本作品策略判断网络流是正常流量与实际情况相吻合的数量，而 TN 是指使用本作品策略判断网络流是异常流量与实际情况相吻合的数量。同时，FP 是指被错误判断成正常流量的异常数据流数量，而 FN 是指被错误判断成异常流量的正常网络流。这些参数用于计算准确率 ACC，该指标可用于衡量和分析性能。由图 4-1 可见，它的 X 轴表示取前几个特征值(即相关性系数由高到低排列的数据特征的数量)，当树的数量为 184 时，取前 7 个特征值，我们的策略效果为最优。因此我们便选取这一组数值作为我们的策略模型的参数（图中数据仅是包含最优数据的部分图表，实际数据所设置树的数量步长为 1，数据规模过大故采用步长为 23 进行绘制）。

根据图示，我们可以很清楚的看出，在本策略中使用的复合特征值筛选算法，最终准确度可达到 93.72%，能够显著保证最终策略检测 DDoS 的准确性。这归功于本作品提出的策略，尤其是在第二层策略中检测网络流数据部分使用的复合特征值筛选方法对特征值的有效筛选，和孤立森林算法对数据进行有效的二元分类。通过复合特征选择算法对 CIC-DDoS2019 数据集进行特征萃取，获得表 4-2 所列的 7 个高关联度特征。该公开数据集高度模拟真实网络环境，所提取的显著性特征对现实网络流量模式识别具有强解释性。

表 4-2 本文方法选取的特征及其功能阐释

特征名词	功能描述
ACK 标志计数	TCP 协议中 ACK（确认）标志位的计数
平均包大小	流中数据包的平均尺寸
子流前向字节数	子流中传输的字节总量
总后向包数	流中每秒传输的数据包数量
空闲均值	数据包之间的平均空闲时间间隔
流包速率	数据包传输速率的计量
包长均值	流中数据包长度的平均值

所选特征体系通过多维度网络流量参数构建了异常检测框架，涵盖多个关键维度。其中，ACK 标志计数用于统计 TCP 数据包中 ACK 位的出现频次，以衡量连接交互的完整性与稳定性；平均包大小与包长均值共同反映流中数据包尺寸分布特征，可用于识别具有集中包长模式的非常规流量；子流前向字节数衡量客户端至服务端的单向数据传输规模，总后向包数则用于评估服务器响应行为，两者结合可识别 DDoS 攻击中常见的通信不对称现象；空闲均值揭示数据包之间的时间间隔，低值通常指示高密度、高频次的恶意访问；流包速率则量化单位时间内的数据包数量，能够捕捉异常突发性流量模式。这些深度关联在网络行为语义上具备高度解释性，共同构建了面向 SDN 环境的动态感知型异常检测基础，为

网络安全态势分析提供了丰富的特征支撑。

4.5.3 DDoS 检测准确性验证

在识别模式和检索信息的领域内，真阳性（TP）、真阴性（TN）、假阳性（FP）和假阴性（FN）这四个二分类指标是用来进行系统性能衡量的核心参数。针对 DDoS 攻击检测研究，我们采用 CIC-DDoS2019 公开数据集的 SYN 子集进行实验，使用 70,336 条正常流量样本进行模型训练，并在包含正常与异常混合流量的 907 条测试集上验证性能。其中 TP 表征正确识别的正常流量，TN 表征正确检测的异常流量，FP 指代正常流量被误判为异常的情况，FN 则反映异常流量被漏检的实例。这些参数被用于计算准确率（ACC）这一关键性能评估指标，其攻击检测准确率计算公式定义如下：

$$\mathrm{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4-18}$$

基于公式(4-18)，我们分别计算并对比了采用皮尔逊相关系数、核方法、距离相关系数、斯皮尔曼相关系数以及 Lasso 正则化技术后，最终 DDoS 检测策略的准确率。通过多维度特征选择算法的性能对比，可评估不同相关性度量方法对攻击检测敏感度的提升效果。皮尔逊相关系数r_{xy}用于度量两个变量x与y间的线性相关关系，其定义为：

$$X_{\mathrm{score}} = \left| \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \right| \tag{4-19}$$

在公式(4-19)中，x_i和y_i表示各样本点，$\overline{x}$和$\overline{y}$为样本均值，n为观测次数。取绝对值是为了确保得分为非负值。核方法（如径向基函数核）通过将数据映射到高维空间以简化分类问题。其 RBF 核定义为：

$$X_{\mathrm{score}} = \sum_{i=1}^{n} \exp\left(-\frac{|x_i - x_i'|^2}{2\sigma^2} \right) \tag{4-20}$$

在公式(4-20)中，x_i和x_i'为数据样本点，σ是控制核函数带宽的关键参数。该方法的评分机制通过计算所有数据点间的核相似度之和实现。距离相关性dCor能够同时捕获变量间的线性和非线性关联关系，其定义为：

$$X_{\mathrm{score}} = \left| \frac{dCov(X,Y)}{\sqrt{dCov(X,X) \cdot dCov(Y,Y)}} \right| \tag{4-21}$$

在公式(4-21)中，$dCov$表征变量X与Y间的距离协方差，用以捕捉二者依赖关系的强度。取绝对值运算确保相关系数保持非负特性。斯皮尔曼相关系数ρ通过评估两变量秩次关系的单调性来度量相关性强度，其计算方式定义为：

$$X_{\text{score}} = \left| 1 - \frac{6\sum d_i^2}{n(n^2-1)} \right| \qquad (4-22)$$

在公式(4-22)中，d_i表示对应样本值的秩次差，n为观测总数。取绝对值运算确保评分保持非负特性。Lasso 正则化方法通过正则化约束实现回归与特征选择的双重目标，其目标函数定义为：

$$X_{\text{score}} = \arg\min_{\beta} \left(\sum_{i=1}^{n} (y_i - \beta_0 - \beta^T x_i)^2 + \lambda|\beta|_1 \right) \qquad (4-23)$$

在公式(4-23)中，y_i为响应变量，x_i为预测变量向量，β表示系数向量，λ是正则化参数，n为观测总数。该评分通过最小化损失函数获得，该函数由残差平方和与L_1范数正则项构成：前者用于拟合数据，后者通过压缩系数绝对值之和来约束模型复杂度并抑制过拟合现象。最终特征系数通过 Lasso 正则化技术得出，其稀疏性特征可实现关键预测变量的自动选择。

表 4-3 不容特征选择算法下的预测准确率对比

特征筛选算法	预测准确率(%)
本文方法	97.66
Lasso 正则化	89.63
Pearson 相关系数	88.74
Spearman 相关系数	87.55
Kernel 方法	86.12
Distance 相关系数	85.22

如表 4-3 所示，本研究所采用的复合特征选择算法作为数据预处理步骤展现出显著优势。该复合特征选择算法最终实现 97.66%的检测准确率，为策略层 DDoS 攻击检测的可靠性提供强保障。这一性能提升得益于本文提出的双层级策略架构：第二层级策略中复合特征值对网络流数据的深度特征筛选，以及孤立森林算法对数据的精准二元分类能力。此外，受文献[60]启发，我们进一步对比了不同特征选择算法的召回率、精确率及 F1 值，以全面评估 DDoS 攻击检测性能。

精确率（Precision）：通过比较本文提出的复合特征选择方法与其他算法的精确率指标，可评估其阳性预测的准确性。精确率衡量在所有被判定为攻击的样本中真实攻击样本的占比，该指标对 DDoS 检测场景尤为重要，因其直接反映算法在避免将正常流量误判为攻击（即降低误报率）方面的能力。具体而言，高精确率意味着安全系统不会因过度敏感而干扰合法网络通信。如图 4-2 所示，各特征选择算法在 DDoS 攻击检测任务中的精确率表现存在显著差异。本文提出的复合特征选择方法达到了 0.9766 的精确率，显著优于其他对比方法，其中 Lasso 方法的精确率为 0.892，Pearson 相关系数方法的精确率为 0.880，Spearman 相关系数方法的精确率为 0.857，Kernel 方法的精确率为 0.823，而距离相关系数方法的精

确率最低，仅为 0.777。这些实验结果表明，本文所采用的复合特征选择策略在攻击检测任务中的表现明显优于其他传统特征选择方法，能够相当大的程度上降低误报率（False Positive Rate, FPR），从而能有效减少正常流量被错误当成 DDoS 攻击流量的概率。

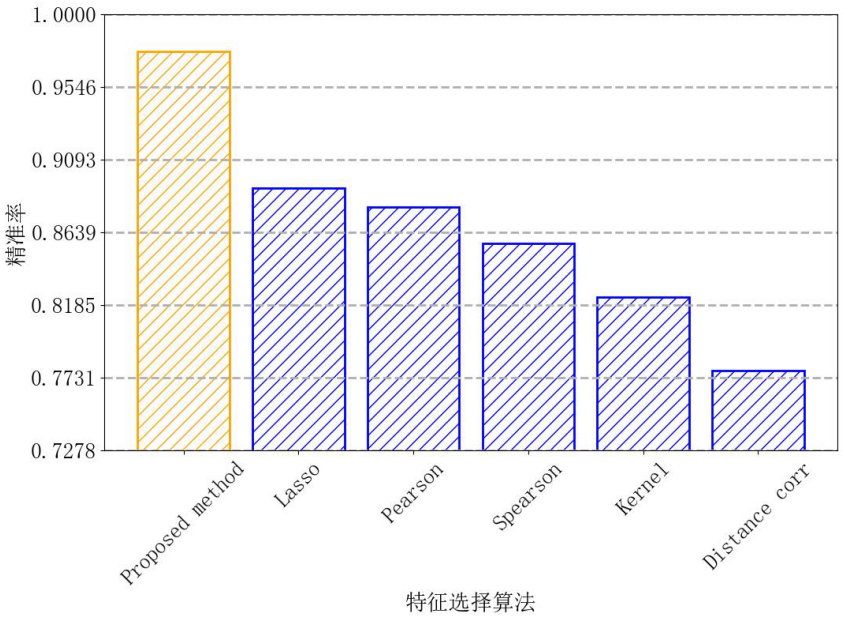

图 4-2 不同特征选择算法的精确率对比分析

召回率：通过对比分析复合特征选择方法与其他算法的召回率，评估其攻击样本的捕获能力。召回率定义为实际攻击样本中被正确识别的比例，是衡量DDoS检测系统漏检风险的核心指标。该指标直接反映算法从真实攻击流量中有效筛选威胁的能力，高召回率意味着防御体系能显著降低攻击突破网络防线的概率。

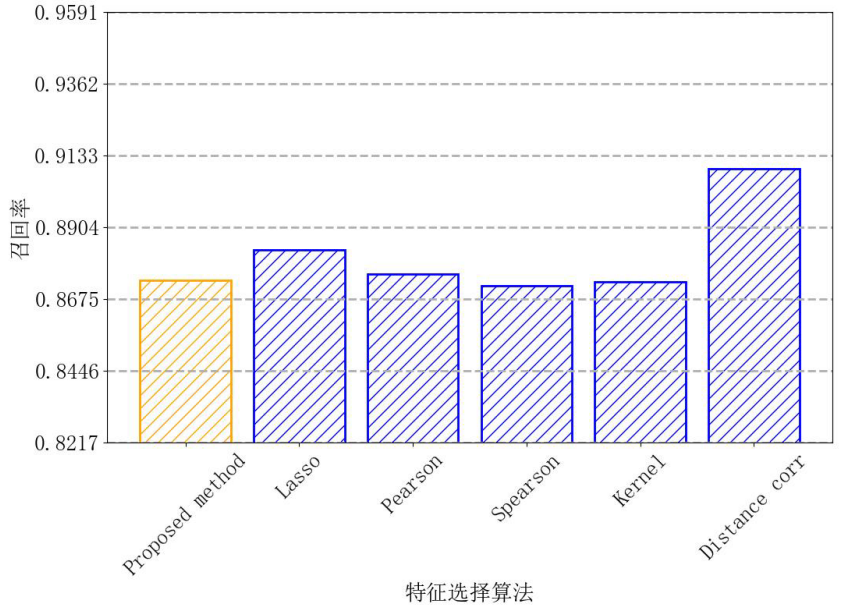

图 4-3 不同特征选择算法的召回率对比评估

如图 4-3 所示各方法召回率表现，本文提出的复合特征选择方法召回率为 0.873，与 Lasso 方法 0.883、Pearson 相关系数 0.875、Spearman 相关系数 0.872、Kernel 方法 0.873 及距离相关系数 0.909 相比虽略有波动，但仍保持对实际攻击样本的强捕获能力。实验结果表明，本方法在真实网络攻防场景中能稳定识别超 87%的 DDoS 攻击流量，体现其设计实用性。

F1 值：最终通过调和平均数综合评估精确率与召回率的均衡性。F1 值作为精确率和召回率的加权调和平均，能有效平衡误报与漏检的权衡关系，这对 DDoS 攻击检测具有特殊意义——既需最大限度捕获真实攻击（高召回），又需避免安全系统因误判干扰正常通信（高精确）。该指标为评估防御系统综合效能提供单一量化标准。

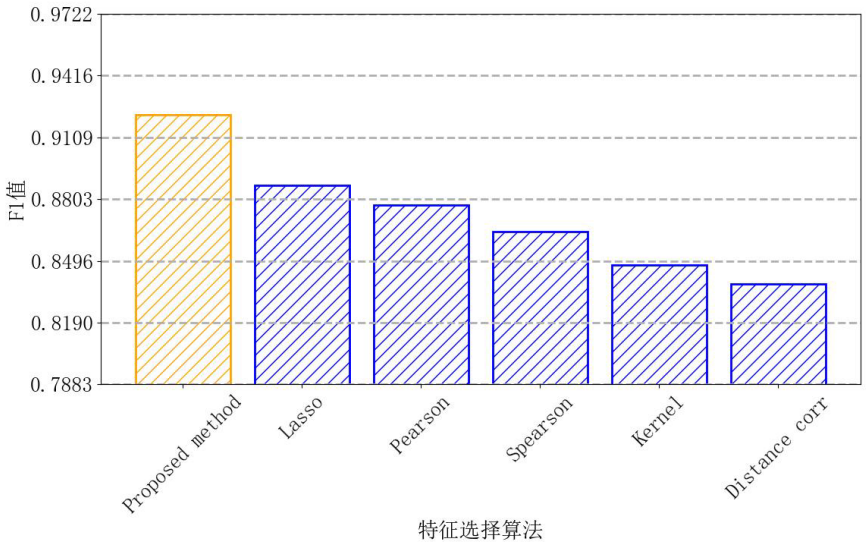

图 4-4 不同特征选择算法的 F1 值对比评估

如图 4-4 所示各方法 F1 值表现，本文提出的复合特征选择方法以 0.922 的 F1 值居首，显著优于 Lasso 方法 0.886、Pearson 相关系数 0.877、Spearman 相关系数 0.865、Kernel 方法 0.846 及距离相关系数 0.838。这一结果表明，本方法在精确率与召回率的均衡性调优方面展现最优性能，即既能通过高精确率减少正常流量的误判干扰，又能通过稳定的召回能力捕获多数真实攻击，最终实现误报率与漏检率的最佳权衡控制。

4.5.4 同类防御策略的异常网络流判定时间比较

基于前文对异常检测特征选择、算法参数确定及检测准确率的系统性阐述，本小节将针对本研究的双层级检测策略与其他同类方案进行异常值分析耗时对比。为评估所提出无监督学习策略的有效性，本文选择当前在 SDN 异常检测中具有代表性的传统方法作为对比基线。具体选择一种机器学习方案和一种深度学

习方案的 DDoS 防御策略作为参照：实时检测方案[25]与增强型 k 均值聚类方案[26]。相较于深度学习模型，这些算法在无需 GPU 支持的情况下即可在普通计算节点上完成训练与推理，计算资源开销相对可控，具有较好的实际部署可行性。与此同时，与基于监督学习的随机森林模型进行对比，也能有效突出无监督策略在无需标签、低开销条件下依然保持较高检测性能的优势。

为确保实验结果公平性，实验统一采用本研究构建的 SDN 网络拓扑环境，并以 CIC-DDoS2019 数据集作为模拟真实流量的数据支撑。在数据预处理阶段，统一应用本文复合特征选择方法计算全部数据特征的相关系数。将数据集划分为训练集与测试集后，对三种策略进行多轮次训练测试。每轮训练均从相关系数最低的特征开始逐步剔除冗余特征。该设计最大限度消除了不同策略特征选择方法差异带来的潜在偏差。最终对比三种策略在相同数量精选特征下完成训练后，对同一测试集进行判定的耗时表现。

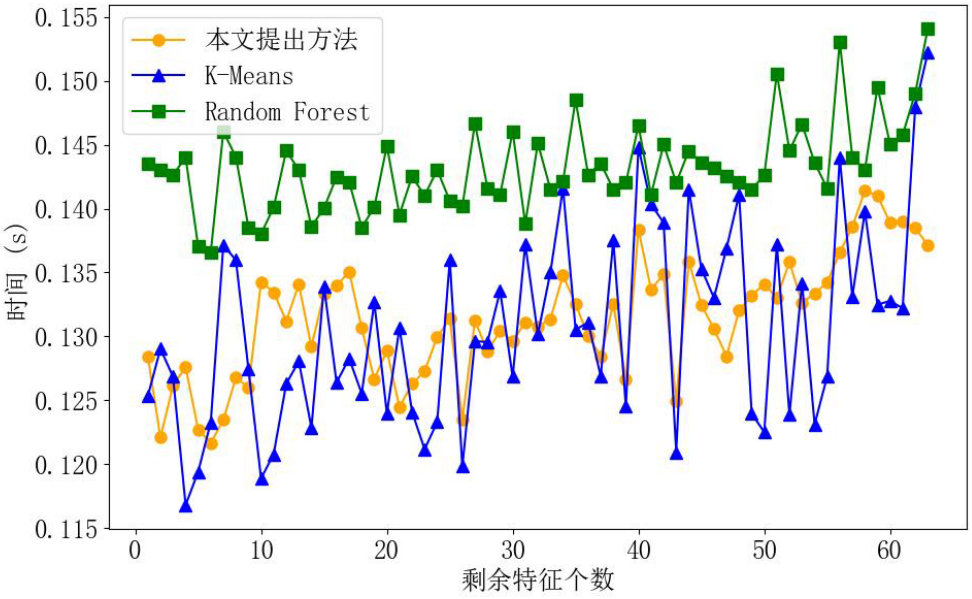

图 4-5 小规模 SDN 架构下 DDoS 防御策略的异常判定耗时对比分析

如图所示，使用随机森林算法的 Real-time detection 策略耗费的时间普遍长于另外两种策略，而我们的双重防御策略同使用 k-means 算法的防御策略时间较为接近，但从上下幅度来看我们的防御策略判定异常时间更为稳定，较少出现大的波动。因此，在判定准确度相近的情况下，我们的策略相较同类型策略所需的判定时间更少，且更为稳定，不易出现大波动。充分体现了我们策略在小型网络拓扑中有较好的防御成效。

4.6 本章小结

本章提出了一种基于区块链和无监督学习技术的双层 DDoS 检测策略，并利

用 CIC-DDoS2019 数据集进行了实验验证。我们首先分析了 DDoS 攻击检测的基本评估指标，包括真阳性（TP）、真阴性（TN）、假阳性（FP）和假阴性（FN），并据此计算攻击检测的准确率（ACC）。

在实验过程中，我们采用 70,336 条正常流量样本进行模型训练，并在包含正常与异常混合流量的 907 条测试集上进行性能评估。实验分别考察了不同特征选择方法（皮尔逊相关系数、核方法、距离相关系数、斯皮尔曼相关系数及 Lasso 正则化技术）对 DDoS 检测准确率的影响。结果表明，不同相关性度量方法在不同场景下具有各自的优势，其中基于核方法的特征选择能有效提升模型的检测能力。实验结果显示，采用双层检测策略后，DDoS 攻击检测的准确率在 CIC-DDoS2019 数据集上达到了 97.66%，相比于传统检测方法提高了 8%。

总之，在这一章中，我们考察了各种特征选择方法在 DDoS 检测策略中的应用效果，已验证了无监督学习双层检测机制是行之有效的。实验结果表明：所提策略能在确保高检测准确率前提下降低误报率与漏报率，是 SDN 背景下 DDoS 防御问题的一种可行解决方案。

5 总结与展望

5.1 总结

本项研究主要聚焦于软件定义网络（SDN）环境中 DDoS 攻击的防御策略，并结合了区块链技术以及无监督学习的方法，提出基于区块链和双层检测机制的流表保护方案，增强 SDN 网络安全性及抗 DDoS 攻击的能力。这两个方面的研究，共同组成了本人基于机器学习和区块链对 SDN 控制器 DDoS 防御能力策略的研究。

首先，对于区块链流表保护，该研究对 SDN 架构下流表管理存在的安全风险进行了分析，并针对流表篡改，伪造以及单点失效的情况提出一种基于区块链存储方案。利用区块链去中心化的特点提升流表完整性与可信度以规避恶意篡改的危险。同时本方案采用智能合约来管理流表，降低人为干预和提高系统自动化程度。在流表转发稳定性方面，通过在 Pox 控制器北向接口集成 FISCO BCOS 区块链组件构建流表存储架构，在相同网络拓扑下对比发现：引入区块链后，数据转发过程中的 Mdev 均值为 0.1122ms，显著低于原始 Pox 控制器的 0.1835ms，延迟降低约 38.86%。该结果验证了所提方案不仅不会带来显著性能负担，反而因智能合约优化与链式存储特性，在一定程度上提升了网络转发的稳定性。在网络拓扑扩展实验中，随着交换机数量从 1 个逐步扩展至 9 个，系统在 1 至 6 个交换机范围内维持良好性能，控制器响应时间与原始 Pox 差异基本控制在 3%以内；当交换机数超过 7 个时，Pingall 平均响应时间出现明显上升，表明当前方案在中小规模网络中具备良好的可扩展性，但在大规模拓扑中存在存储与共识延迟瓶颈，需在后续研究中引入轻量链或分片机制进一步优化。

针对 DDoS 攻击检测问题，研究提出基于无监督学习并将复合数据特征筛选与机器学习模型相结合的双层检测策略来进行有效攻击检测。第一层使用轻量级的检测方法快速筛选，降低误报率，提高检测效率；在第二层，我们采用了复合特征筛选技术和更为细致的无监督学习模型，对可疑的流量进行了深入的分析，从而增强了检测的准确率。在 DDoS 攻击检测实验中，首先通过复合特征选择机制对原始 77 维统计特征进行降维处理，综合互信息（MI）、随机森林（RF）与递归特征消除（RFE）算法，筛选出 7 个高区分度的关键特征，包括 ACK 标志计数、空闲均值、平均包大小、子流前向字节数、总后向包数、包长均值与流包速率。这些特征深度结合 SDN 中典型异常流行为特征，有效提升了模型的表达能力与泛化性能。最终实验显示，所提策略在测试集上检测准确率达 97.66%，同时在维持高准确率前提下，有效降低误报率与漏报率，证明该方案在多种攻击

场景下具有良好的鲁棒性与适应性。

通过对上述过程的研究与验证得出，本文所设计的防御方案对增强 SDN 网络的安全性有明显的效果。在保障流表安全、提升检测性能方面具有突出优势，在 SDN 网络面向未来规模化部署的场景下具备良好的可行性与应用前景。

5.2 展望

尽管本研究在 SDN 安全防御方面取得了一定的成果，但仍然存在一些值得进一步探索的方向：

1、优化流表保护方面，针对大型 SDN 网络拓扑对区块链存储与计算开销需求大，今后可与轻量级区块链或者分片技术相结合，在流表管理方面做进一步优化，满足大规模网络环境下使用。

2、对检测模型进行改进，无监督学习方法对攻击检测表现出良好的适应性，但是仍然存在模型精度与实时性的取舍问题。今后可将联邦学习或者强化学习相结合以增强检测模型泛化能力与实时响应能力。

3、防御策略加强：DDoS 攻击探测和防御是本次研究的重点，今后可进一步拓展至流量投毒攻击和旁道攻击这类更为宽泛的安全威胁中去，从而建立更为综合的 SDN 安全体系。

4、系统部署及实测：本次研究实验以仿真环境为主，今后可以对实际 SDN 的部署进行较大范围的试验，以评价该方案在现实网络环境下的可行性及鲁棒性，对该方案工程化实施做进一步的优化。

综上，该研究对 SDN 环境中 DDoS 攻击防御问题提出了创新解决思路，验证了该方法的有效性。今后的研究可以着眼于存储优化，检测改进以及系统部署，进一步提高 SDN 网络安全性与可扩展性。

参考文献

[1] 第 55 次《中国互联网络发展状况统计报告》发布[J].传媒论坛,2025,9(02):121.

[2] Nisar K, Jimson E R, Hijazi M H A, et al. A survey on the architecture, application, and security of software defined networking: Challenges and open issues[J]. Internet of Things, 2020, 12: 100289.

[3] Luo S, Wu J, Li J, et al. A defense mechanism for distributed denial of service attack in software-defined networks[C]//2015 Ninth International Conference on Frontier of Computer Science and Technology. IEEE, 2015: 325-329.

[4] McKeown N, Anderson T, Balakrishnan H, et al. OpenFlow: enabling innovation in campus networks[J]. ACM SIGCOMM computer communication review, 2008, 38(2): 69-74.

[5] Kaur P, Kumar M, Bhandari A. A review of detection approaches for distributed denial of service attacks[J]. Systems Science & Control Engineering, 2017, 5(1): 301-320.

[6] Ahmad Z, Shahid Khan A, Wai Shiang C, et al. Network intrusion detection system: A systematic study of machine learning and deep learning approaches[J]. Transactions on Emerging Telecommunications Technologies, 2021, 32(1): e4150.

[7] 赵凯琳, 靳小龙, 王元卓. 小样本学习研究综述[J]. 软件学报, 2020, 32(2): 349-369.

[8] Breiman L. Random forests[J]. Machine learning, 2001, 45: 5-32.

[9] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.

[10] Sutrala A K, Obaidat M S, Saha S, et al. Authenticated key agreement scheme with user anonymity and untraceability for 5G-enabled softwarized industrial cyber-physical systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(3): 2316-2330.

[11] Kaur S, Sandhu A K, Bhandari A. Investigation of application layer DDoS attacks in legacy and software-defined networks: A comprehensive review[J]. International Journal of Information Security, 2023, 22(6): 1949-1988.

[12] Latah M, Kalkan K. When SDN and blockchain shake hands[J]. Communications of the ACM, 2022, 65(9): 68-78.

[13] Saha S, Chattaraj D, Bera B, et al. Consortium blockchain-enabled access control mechanism in edge computing based generic Internet of Things environment[J]. Transactions on Emerging Telecommunications Technologies, 2021, 32(6): e3995.

[14] Abou El Houda Z, Hafid A, Khoukhi L. Co-IoT: A collaborative DDoS mitigation scheme in IoT environment based on blockchain using SDN[C]//2019 IEEE Global Communications Conference (GLOBECOM). IEEE, 2019: 1-6.

[15] Chattaraj D, Saha S, Bera B, et al. On the design of blockchain-based access control scheme

for software defined networks[C]//IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2020: 237-242.

[16] Abuhashim A, Tan C C. Smart contract designs on blockchain applications[C]//2020 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2020: 1-4.

[17] Giotis K, Apostolaki M, Maglaris V. A reputation-based collaborative schema for the mitigation of distributed attacks in SDN domains[C]//NOMS 2016-2016 IEEE/IFIP network operations and management symposium. IEEE, 2016: 495-501.

[18] Wani S, Imthiyas M, Almohamedh H, et al. Distributed denial of service (DDoS) mitigation using blockchain—A comprehensive insight[J]. Symmetry, 2021, 13(2): 227.

[19] Valdovinos I A, Pérez-Díaz J A, Choo K K R, et al. Emerging DDoS attack detection and mitigation strategies in software-defined networks: Taxonomy, challenges and future directions[J]. Journal of Network and Computer Applications, 2021, 187: 103093.

[20] Gao Y, Chen Y, Hu X, et al. Blockchain based IIoT data sharing framework for SDN-enabled pervasive edge computing[J]. IEEE Transactions on Industrial Informatics, 2020, 17(7): 5041-5049.

[21] Núñez-Gómez C, Carrión C, Caminero B, et al. S-HIDRA: A blockchain and SDN domain-based architecture to orchestrate fog computing environments[J]. Computer Networks, 2023, 221: 109512.

[22] Zeng Z, Zhang X, Xia Z. Intelligent blockchain-based secure routing for multidomain SDN-enabled IoT networks[J]. Wireless Communications and Mobile Computing, 2022, 2022(1): 5693962.

[23] Huang X, Zheng K, Chen S, et al. Construction of switch information security protect ion system based on software-defined networking[J]. Transactions on Emerging Teleco mmunications Technologies, 2024, 35(9): e5033.

[24] 任瑞峰.基于区块链的软件定义物联网多控制器安全研究[D].内蒙古科技大学,2023.DOI:10.27724/d.cnki.gnmgk.2023.000611.

[25] Ma R, Wang Q, Bu X, et al. Real-Time Detection of DDoS Attacks Based on Random Forest in SDN[J]. Applied Sciences, 2023, 13(13): 7872.

[26] Marvi M, Arfeen A, Uddin R. An augmented K-means clustering approach for the detection of distributed denial-of-service attacks[J]. International Journal of Network Management, 2021, 31(6): e2160.

[27] 万梅,曹琳.基于强化学习的软件定义网络安全[J].计算机工程与设计,2020,41(08):2128-2134.DOI:10.16208/j.issn1000-7024.2020.08.006.

[28] 杨桂芹,张蔚,张若.SDN 环境下基于 Rényi-RF-XGBoost 的 DDoS 攻击检测研究[J].兰州交通大学学报,2025,44(01):28-38.

[29] 王坤,付钰,段雪源,等.基于深度学习的 SDN 异常流量分布式检测方法[J].通信学报,2024,45(11):114-130.

[30] Lian W, Li Z, Guo C, et al. FRChain: a blockchain-based flow-rules-oriented data forwarding security scheme in SDN[J]. KSII Transactions on Internet and Information Systems (TIIS), 2021, 15(1): 264-284.

[31] Su J, Jiang M. A hybrid entropy and blockchain approach for network security defense in SDN-based IIoT[J]. Chinese Journal of Electronics, 2023, 32(3): 531-541.

[32] Li W, Wang Y, Meng W, et al. BlockCSDN: towards blockchain-based collaborative intrusion detection in software defined networking[J]. IEICE TRANSACTIONS on Information and Systems, 2022, 105(2): 272-279.

[33] 宋金阳.区块链环境下基于机器学习的 DDoS 攻击检测方法研究[D].海南大学,2022.DOI:10.27073/d.cnki.ghadu.2022.000702.

[34] 杨昊.基于公有链的 DDoS 攻击轻量级检测方法研究[D].海南大学,2022.DOI:10.27073/d.cnki.ghadu.2022.000262.

[35] Keshari S K, Kansal V, Kumar S. A systematic review of quality of services (QoS) in software defined networking (SDN)[J]. Wireless Personal Communications, 2021, 116(3): 2593-2614.

[36] Zheng S. Research on SDN-based IoT security architecture model[C]//2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). IEEE, 2019: 575-579.

[37] 张健,朱丹.软件定义网络 DDoS 攻击问题研究[J].信息通信技术与政策,2025,51(01):56-63.

[38] Lukaseder T, Stölzle K, Kleber S, et al. An sdn-based approach for defending against reflective ddos attacks[C]//2018 IEEE 43rd Conference on Local Computer Networks (LCN). IEEE, 2018: 299-302.

[39] Nugraha B, Murthy R N. Deep learning-based slow DDoS attack detection in SDN-based networks[C]//2020 IEEE conference on network function virtualization and software defined networks (NFV-SDN). IEEE, 2020: 51-56.

[40] 沈瑞.区块链共识算法的研究与应用[D].南京邮电大学,2021.DOI:10.27251/d.cnki.gnjdc.2021.000197.

[41] 任博,邱国栋.克服合谋掩饰行为：智能区块链与供应链金融运行机制耦合[J].中国流通经济,2022,36(03):35-47.DOI:10.14089/j.cnki.cn11-3664/f.2022.03.004.

[42] Breunig M, Kriegel H P, Ng R T, et al. LOF: identifying density-based local outliers [C]//Proceedings of the 2000 ACM SIGMOD international conference on Management of data. 2000: 93-104.

[43] Schubert E, Sander J, Ester M, et al. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN[J]. ACM Transactions on Database Systems (TODS), 2017, 42(3): 1-21.

[44] Jolliffe I T, Cadima J. Principal component analysis: a review and recent developments[J]. Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, 2016, 374(2065): 20150202.

[45] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]//Proceedings of the 25th international conference on Machine learning. 2008: 1096-1103.

[46] Said Elsayed M, Le-Khac N A, Dev S, et al. Network anomaly detection using LSTM based autoencoder[C]//Proceedings of the 16th ACM symposium on QoS and security for wireless and mobile networks. 2020: 37-45.

[47] 刘宇啸, 陈伟, 张天月. 基于稀疏自动编码器的可解释性异常流量检测[J]. 信息网络安全, 2023, 23(7): 74-85.

[48] Liu F T, Ting K M, Zhou Z H. Isolation forest[C]//2008 eighth ieee international conference on data mining. IEEE, 2008: 413-422.

[49] Wang R, Ye K, Meng T, et al. Performance evaluation on blockchain systems: a case study on Ethereum, Fabric, Sawtooth and Fisco-Bcos[C]//Services Computing–SCC 2020: 17th International Conference, Held as Part of the Services Conference Federation, SCF 2020, Honolulu, HI, USA, September 18–20, 2020, Proceedings 17. Springer International Publishing, 2020: 120-134.

[50] Zheng Z, Xie S, Dai H N, et al. An overview on smart contracts: Challenges, advances and platforms[J]. Future Generation Computer Systems, 2020, 105: 475-491.

[51] De Oliveira R L S, Schweitzer C M, Shinoda A A, et al. Using mininet for emulation and prototyping software-defined networks[C]//2014 IEEE Colombian conference on communications and computing (COLCOM). Ieee, 2014: 1-6.

[52] Tripathy D N, Reed W M. Pox[J]. Diseases of poultry, 2013: 333-349.

[53] Cano-Benito J, Cimmino A, García-Castro R. Toward the ontological modeling of smart contracts: a solidity use case[J]. IEEE Access, 2021, 9: 140156-140172.

[54] Wang S, Balarezo J F, Chavez K G, et al. Detecting flooding DDoS attacks in software defined networks using supervised learning techniques[J]. Engineering science and technology, an international journal, 2022, 35: 101176.

[55] Sharafaldin I, Lashkari A H, Hakak S, et al. Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy[C]//2019 international carnahan conference on security technology (ICCST). IEEE, 2019: 1-8.

[56] Hariri S, Kind M C, Brunner R J. Extended isolation forest[J]. IEEE transactions on knowledge and data engineering, 2019, 33(4): 1479-1489.

[57] Estévez P A, Tesmer M, Perez C A, et al. Normalized mutual information feature selection[J].

IEEE Transactions on neural networks, 2009, 20(2): 189-201.

[58] Biau G, Scornet E. A random forest guided tour[J]. Test, 2016, 25: 197-227.

[59] Chen X, Jeong J C. Enhanced recursive feature elimination[C]//Sixth international conference on machine learning and applications (ICMLA 2007). IEEE, 2007: 429-435.

[60] Yacouby R, Axman D. Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models[C]//Proceedings of the first workshop on evaluation and comparison of NLP systems. 2020: 79-91.

攻读学位期间的学术论文与研究成果

[1] **Tian J**, Shu Z, Chen S, et al. DDoS Defense Strategy Based on Blockchain and Unsupervised Learning Techniques in SDN[J]**.** IEEE Transactions on Network and Service Management, 2024.（第一作者，在投）

[2] **Tian J**, Shu Z, Chen S, et al. Enhanced DDoS Defense in SDN: Double-Layered Strategy with Blockchain Integration[C]//2024 13th International Conference on Communications, Circuits and Systems (ICCCAS). IEEE, 2024: 380-384.（第一作者，已出版）

[3] 王媛滔,舒兆港,钟一文等. 基于 VNF 实例共享的服务功能链部署算法[J].计算机应用研究, 2023, 40(6).（第五作者，已出版）

[4] 邱彩钰,舒兆港,陈淑武等. 基于模糊层次分析和 DQN 的流转发路径探索算法[J].计算机工程与设计,2025.（第六作者，已录用）

[5] 获得奖项：学业一等奖学金

[6] 获得奖项：全国高校计算机能力挑战赛 c++组省赛（三等奖）

[7] 获得奖项：第十四届 MathorCup 数学应用挑战赛华东赛区（三等奖）

[8] 获得奖项：2024 中国高校计算机大赛-网络技术挑战赛省赛（三等奖）

[9] 获得奖项：第十五届蓝桥杯大赛软件赛省赛（C++ 组三等奖）

[10] 获得奖项：学业二等奖学金

[11] 获得奖项：“华为杯”第二十届中国研究生数学建模竞赛（三等奖）

[12] 获得奖项：校数学建模大赛（一等奖）

[13] 获得奖项：校研究生程序设计大赛（三等奖）

[14] 获得奖项：2023 中国高校计算机大赛-网络技术挑战赛省赛（三等奖）

致 谢

三年的读研生涯即将步入尾声，回首自己这期间的点滴，有过迷茫，有过受挫，也有过意气风发下的飘飘然，不过起伏间更多感到的是充实。无论是为了论文反复推倒重来的实验，还是为了下一阶段焦头烂额的专业知识补足。几年间，不知不觉竟也完成了些自己以前一直想做的事。

在此，首先要感谢舒兆港老师带我迈入了 SDN 研究的大门。这一领域的分支庞大，极大的拓展了我的视野，接触到了区块链，机器学习，深度学习等较为前沿的技术。期间老师的指导与建议，也让我在关键时刻不断明晰了前进方向。其次我要感谢浩贤师兄，媛滔师姐，治方，智伟师兄。几位师兄师姐平日的指点与帮助，不仅让我更快的融入到实验室的日常研究中，也给我的校园生活增添了很多欢乐。然后，感谢我的同门彩钰，结下的“战友”情谊，务须多言。还要感谢人数愈发壮大的师弟们。大伙面对问题各显神通，我亦受益良多。当然，也少不了感谢我的几位好室友，千里投缘，珍贵难得。此外，感谢自己的父母，是他们的支持让我在读研过程中能够不断前行。浮云朝露，缘分可遇不可求，也希望自己有幸收获他人给予的同时，也能给身边的大家带来哪怕一点帮助，如若有此，不负此行。

文至末尾，再次回顾三年前的我，彼时虽左支右绌，可仍给自己指了条雾霭朦胧的路，一路走来，至此回头，颔首致意。